Front. Eng. Manag. 2022, 9(2): 337–342
https://doi.org/10.1007/s42524-022-0189-1

COMMENTS

Jianxun YANG, Wenjing WU, Miaomiao LIU, Jun BI

Leveraging individual-level data to advance air pollution health risk management

1 Introduction

Environmental health risk management is a systematic engineering task, engaging multiple disciplines from the academic and government sectors. Reducing environmental health risks has become one of the key targets in the United Nations Sustainable Development Goals (SDGs). This target has been translated into public policies at many jurisdictional levels (Yue et al., 2020). To design region-specific and targeted policy initiatives, understanding how environmental health risks are spatially distributed and temporally resolved is fundamental.

Along with the advances in high-resolution pollution mapping and projection, environmental risk assessment and management have been performed at a highly granular scale (Caplin et al., 2019). For instance, taking advantage of the fine air pollution dataset, recent research efforts have evaluated the health burdens of historical air pollution exposure or cost-benefits of pollution control policies at finer-scale administrative units and grids (Liu et al., 2017; Ou et al., 2020). These assessments reveal expensive health costs of air pollution exposure and highlight that the health outcomes of air pollution are unevenly distributed across regions and more evident among vulnerable populations (Colmer et al., 2020).

Even though detailed geospatial mappings of air pollution capture local patterns, they do not necessarily represent individual-level unique exposure experiences and health outcomes. Personal exposure to air pollution can be influenced by a range of behavioral factors such as mobility patterns and self-protective actions (Tainio et al., 2021). Individual cofactors such as risk perceptions of air pollutants, baseline health conditions, and socio–economic status also influence the pollution-related health impacts (Piel et al., 2020). Dramatic variations in health risks, therefore, occur even within small spatial units such as blocks and neighborhoods. Thus, individual-level exposure and health assessment are critical in advancing the engineering management of air pollution at granular spatiotemporal scales and informing targeted local policies for reducing pollution-related health risks.

To perform air pollution health risk assessment and management at the individual level, first, gathering datasets relevant to personal exposure experiences and health outcomes is important. For example, portable sensor technology is widely applied, in which participants carry the sensors to measure their real-time locations and micro-environmental exposure (Su et al., 2017). These monitors can illustrate individual time–activity patterns and assess personal-specific pollution exposure during outdoor activities such as daily commuting and exercises (Dons et al., 2017). These activities finally form direct flows of datasets on human behaviors and are made available to researchers for more nuanced characterizations of environmental risk.

In this comment, we summarize different types of individual-level data and outline pathways through which the data may advance air pollution health risks assessment. We then review representative studies revolving around these aspects and showcase how abundant information at the individual level improves environmental health risk management. We finally detail the challenges and uncertainties in this rapidly growing field and highlight the priorities in future research. The aim is to motivate local policy actions and foster collective research efforts to promote public health.

Received November 24, 2021; accepted December 24, 2021

Jianxun YANG, Wenjing WU, Miaomiao LIU (✉), Jun BI
State Key Laboratory of Pollution Control and Resource Reuse, School of Environment, Nanjing University, Nanjing 210023, China
E-mail: liumm@nju.edu.cn

This work is supported by the National Natural Science Foundation of China (Grant Nos. 71921003, 72174084, and 71761147002) and the Fundamental Research Funds for the Central Universities (Grant No. 0211-14380171).

2 Overview of individual-level data

Before diving into specific aspects to which individual-level data may contribute, we first overview major types of individual-level data and common approaches to data acquisition, as well as their strengths and weaknesses. We define individual-level data as those reflecting spatial and behavioral patterns, socio–economic and demographic covariates, and mental or physiological health markers related to individuals. Traditional individual-level datasets used in previous studies are often seen in large follow-up surveys (Hanigan et al., 2019). These studies investigated the influence of sociodemographic cofactors (e.g., age, gender, and income) and potential confounders such as smoking on the health impacts of ambient pollution. However, personal heterogeneities in risk perception, short-term mobility, and behavioral patterns are often ill-represented in cohort studies. In recent years, rapid development in cheap monitoring devices and data extraction techniques has enabled researchers to understand individuals' perceptions, behaviors, and health outcomes with smaller costs (Chaix, 2018). We focus on the datasets built upon the recent methodological progress. Here, we identify two categories of individual-level data that can be used to advance air pollution health risk management (Fig. 1).

2.1 Participatory sensing data

The first category of individual-level data gathers information with the involvement of the person being observed. These data-collection efforts can involve the voluntary participation of individuals aware of the research purpose and carrying devices for data recording. Typical examples include data collected from time–activity surveys or diaries, low-cost mobile sensing, and lab-based experiment. These datasets have relatively small sample sizes from dozens to thousands of participants. Time–activity diaries record how much time a person has spent on specific activities in different types of locations. They are often made via self-reported questionnaires and have been extensively used as proxies to measure individual physical activities, travel modes, and visited places. The diary-based method has the limitation that activities not in the checklists or of short duration are often underestimated (Vanroy et al., 2014).

As a means of improving personal monitoring, the development of cheap multi-sensor devices has recently enabled researchers to record participants' geo-coordinates, micro-environmental pollution exposure, and physiological status. To determine personal exposure, time–activity data logged by smartphones or wearable bracelets are paired with the timestamped pollution concentrations recorded by portable sensors. Although the data gained by personal monitoring have a relatively high degree of information detail, they somehow are limited by small spatial coverage and short periods. The data quality also largely relies on the measurement accuracy, stability, and durability of sensors (Deville Cavellin et al., 2016).

Finally, experimental human exposure to air pollutants, conducted in a laboratory setting, has been essential to measure associated adverse health effects and reveal potential mechanisms. The technique is now available for

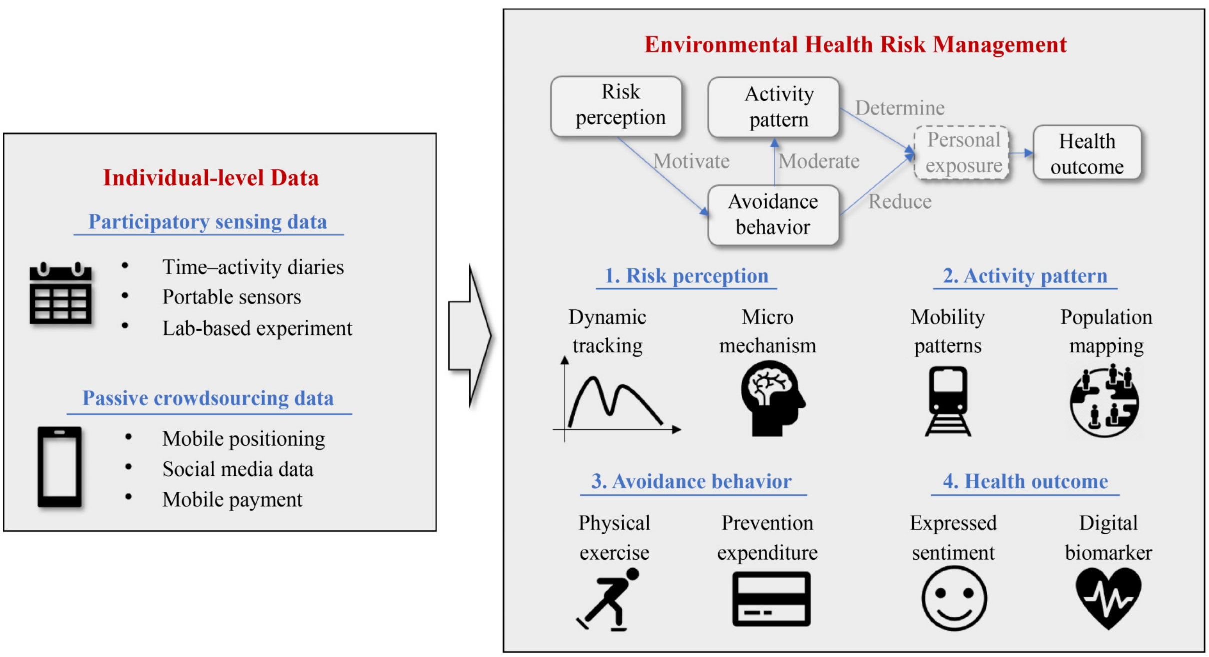

Fig. 1 Applications of individual-level data in environmental health risk management.

generating controllable concentrations of pollutants such as ozone and $PM_{2.5}$ in small chambers. The main advantage of exposure experiments lies in their ability to effectively control the confounding factors and reflect the exposure–response relationship at the individual levels. However, due to limited sample sizes and the challenges of medical ethics, this research design may fail to fully reflect a broad spectrum of populations.

2.2 Passive crowdsourcing data

The second emerging category of individual-level data is often retrieved via passive data collection. That is, citizens generate and upload data continuously without original intention for research purposes. The dramatic increase in passive data volume results from the rise in smart electronic devices and the consequent changes in social activities. These devices continually collect individual data and provide near-real-time insight into human mobility, mental state, and patterns of consumption or communication. Compared with traditional ways of data collection, passive crowdsourcing has a significantly larger size of data samples and is less labor-intensive with a lower cost.

For example, mobile phones provide a wealth of positioning information for individual users, which can be used to estimate the mobility of citizens and map the dynamics of the population (Deville et al., 2014). Social media is an emerging tool for gauging social consensus, where a huge pool of people receive news and vocalize opinions, along with individual-level interaction networks, geotags, timestamps, and text logs for each user (Ghermandi and Sinclair, 2019). The rich information on social media platforms can measure a person's sentiment dynamics at specific locations. Last but not least, the consumer spending datasets, represented by UnionPay consumption and mobile payment data, can be deployed to track personal healthcare expenditure.

The major challenges of individual-level data collected from passive crowdsourcing lie in two dimensions. First, to protect personal privacy in data acquisition and use, some demographic details such as gender, age, and socio–economic status are not accessible or not real. This shortcoming is closely linked to the second concern, namely, the potential bias in the representativeness of the sampled users. For example, social media platforms users tend to be younger people, and data are rich in populous urban areas with supporting infrastructure.

3 Applications of individual-level data in air pollution health risk management

In this section, we introduce the conceptual framework that depicts how individual-level data can advance environmental health risk management, as illustrated in Fig. 1. Health outcomes of pollutants as the endpoints of risk management are determined by personal exposures. Personal exposures are closely linked to individual-level activity patterns. Avoidance behaviors motivated by risk perceptions, including all types of prevention measures, can help moderate activity patterns or directly reduce pollutant exposure. Accordingly, we summarize four aspects for the applications of individual-level data: Tracking dynamics and revealing mechanism of risk perceptions, measuring personal avoidance behaviors, delineating or predicting spatial and temporal activities, and enriching the measurements of health outcomes.

3.1 Risk perception

Prior research has demonstrated that risk perception influences a person's behavioral response to environmental risks. Individuals who underestimate environmental risks are rendered vulnerable because they are reluctant to take the necessary steps for self-protection (Hong et al., 2021). Furthermore, environmental risk perception is dynamic and cyclical, influenced by sudden risk events, mass media news, and other sociocultural factors, which are highly localized. Continuous monitoring of the dynamic of public risk perception is thus crucial.

The psychometric paradigm is a well-established technique for assessing risk perception, which relies on large-scale questionnaire surveys. However, the method is confined to the time-intensive and expensive survey process and thus destined to use small samples and report at hardly operational coarser spatial scales. Individual-level data collected from social media platforms can help address the research gap. Topic modeling and other natural language processing algorithms can analyze the content of big text data published on social media and track the dynamics of public perception of environmental risks. For instance, Zheng et al. (2019) evaluated the public perception of air pollution using data from 400 million microblogs. They found that when the pollution level increased, the emotions expressed on social media decreased. A study investigating the Internet's amplification effect found that sentiments embodied in online information contribute to the increase in air pollution risk perception (Guo and Li, 2018).

Objective and perceived environmental risk levels are mismatched. Risk perception bias reduces social well-being and should be considered when designing risk communication strategies. Individual-level data collected from the lab-based experiment can provide micro mechanism explanations of risk perception bias. Yang et al. (2021) designed a series of psychophysiological studies that examined individuals' emotional responses to landscape photos with different levels of $PM_{2.5}$ concentrations. They validated the Peak-End Rule and Stress Recovery Theory that can explain individual perception bias towards air pollution. Further on, applying an event-related potential and other neuro-scientific techniques, Qin and Han (2009) revealed the brain activity mechanism corresponding to individual perceptions of different types of environmental risks. By fully using these interdisciplinary techniques, researchers can discover targeted intervention strategies for fostering accurate environmental risk perceptions among the general public.

3.2 Activity pattern

Many developing countries like China are undergoing rapid urbanization, leading to frequent population migration between urban and rural areas, continuous expansion of urban space scale, and more active and complex migration patterns within cities. Drastic changes in individual physical activity patterns alter the time and frequency of exposure to environmental risk factors.

Traditional environmental risk assessment technology assumes that individuals have the same exposure behavior pattern, and relies on annual census data or grid estimation data, resulting in low spatial and temporal accuracy of assessment results. These relative "static" estimates fail to account for time spent in different outdoor and indoor micro-environments. Portable and low-cost sensors help fill this gap because they allow for measurements of real-time exposures to airborne pollutants at the individual level across urban locations. This technique has a wide range of application scenarios, with the exposure measurement of commuters being the most typical. For example, a study in Minneapolis examined pollution exposure at different stages of daily commuting and found that the road sections requiring high levels of physical activity also had very high concentrations of pollutants (Hankey et al., 2017). Data collected from time–activity diaries also help model individual heterogeneity in daily mobility. In another study, the authors built a behavior model based on 90000 survey data obtained from residents in Hong Kong. They found that working adults and students have higher exposures because of their higher mobility (Tang et al., 2018).

Passive crowdsourced datasets, such as mobile positioning signals, have enabled fast and cheap production of population maps for evaluating population-weighted exposure to air pollution. Dewulf et al. (2016) simulated the NO_2 exposure of the Belgian population based on the data of five million anonymous mobile phone users. They found that the traditional assessment method ignoring individual exposure behavior would underestimate the overall exposure. Nyhan et al. (2016) used the location information of mobile phone users in New York City for four months, combined with $PM_{2.5}$ concentration data, to quantify the population-weighted $PM_{2.5}$ exposure health risks. Their results suggested significant differences in exposure levels in different time ranges. Integrating spatio–temporal positioning data from participatory sensing or passive crowdsourcing presents opportunities to improve the characterization of individual exposure behavior and the understanding of heterogeneities in population exposure.

3.3 Avoidance behavior

As the environmental risk perception level of the general public increases unceasingly, a large segment of the population has adopted so-called avoidance behaviors to lower the adverse health effects of pollution. Such activities include buying air purifiers and masks, changing travel routes, purchasing green products, and increasing health care expenditure. An on-site survey experiment for 2000 travelers showed that people make sub-optimal, overly risk-averse choices by reducing active commuting (Fan et al., 2021). Pollution-avoidance behaviors may help reduce the health impact of environmental risks, but these measures' effectiveness has been rarely evaluated, and appropriate measurements remain lacking.

Avoidance behaviors of individuals have not yet been systematically included in the risk assessment framework. Previous studies have measured subjective adaptive motivations, such as willingness to pay through questionnaire surveys, but failed to trace actual adaptive behaviors. Ignoring individual avoidance behavior may amplify environmental inequality because disadvantaged groups are less likely to spend extra money on pollution-avoidance strategies (Sun et al., 2017).

With the prosperity of the Internet economy, massive payment data generated by consumers, such as mobile payments, can address existing gaps and provide a measurement of avoidance behavior. Numerous types of consumer spending can be used to indicate avoidance behavior, such as health insurance claims, hospital and pharmacy visits, and medical records on purchasing preventive pharmaceuticals. For example, using the data of credit and debit card transactions in China from 2013 to 2015, Barwick et al. (2018) found that the lowest income group had the largest increase in household healthcare expenditure after exposure to $PM_{2.5}$. By deeply mining the individual-level consumption data and tracking the expenditure of goods or services related to avoidance behavior, we can identify vulnerable groups who are truly at risk.

3.4 Health outcome

Emerging individual-level data are also associated with more feasible options for measuring health outcomes. Some of the measurements might not be new for certain disciplines, but their integration with pollution exposure has increased our understanding of the consequence of environmental risk. They also help to draw a more comprehensive picture to inform the social benefits of engineering measures of risk reduction.

The negative sentiments expressed on social media by individuals reflect their poor mental health states. These psychological states are linked to low work productivity, depression, and, at worst, increasing rates of suicide. Therefore, social media platforms have provided low-cost large-cohort datasets to monitor public psychological wellbeing. Using 1 billion Twitter updates, for instance, Baylis (2020) calculated the sentiment score for each tweet and estimated the relationship between expressed sentiment on Twittersphere and outdoor ambient temperature. He demonstrated significant declines in online sentiment from hot to cold temperatures.

In other application scenarios, some mild health outcomes such as pharmacy visits, respiratory medication sales, and cardiorespiratory drug prescriptions can be retrieved from individual-level data like mobile positioning data and payment records (Casas et al., 2016). These mild adverse health impacts of pollution exposure have induced enormous social costs but received little attention in risk assessment studies. Portable sensors have enabled researchers to measure physical indicators for health monitoring and diagnoses, such as heart rate variability, sleep patterns, blood pressure, and self-tracking surveys (Tsou et al., 2021). In lab-based experiments, health metrics from medical science such as biomarkers of inflammation and physiological signals such as electroencephalograms can be applied to measure the acute health impacts of pollution exposure on physical health and cognitive ability.

4 Outlooks for future research

Given that human and environmental systems are highly networked, we have seen a growing body of studies leveraging individual-level data to advance environmental health risk assessment and management. In a new age of environmental research, we highlight a few potential directions fostering the utilization of individual-level data in systematic engineering.

First, the procedure of data management and quality assurance for analyzing high-resolution pollution exposure should be improved. Given that the individual-level data come from a wide range of sources and scales, challenges are prominent when integrating various datasets regarding ethical issues, long-term data availability, noise, and bias corrections. Therefore, a framework synthesizing different types of data is urgent and crucial to improve data utilization efficiency, whether for academicians in risk analysis, environmental engineers, or local policymakers.

Second, the integration of multidisciplinary methods should be promoted. Despite recent progress, studies in this field are scattered across various disciplines, including environmental epidemiology, physiology and psychology, behavioral science, social and economic science, geographic information systems, and big data science. Investigating approaches to method integration can help identify priority areas for future research and engineering practice.

Third, data applications should be aligned with national strategies and United Nations SDGs. For example, in the Chinese context, the government has pledged to peak its carbon emissions by 2030 and achieve carbon neutrality by 2060. In the coming decades, challenges from ambitious carbon reduction and pollution elimination plans will coexist, calling for synergetic management of multiple environmental risks in a granular fashion. As advocated in the SDG frameworks, the commitment to protect environmental health and enhance regional equality has also informed a research agenda for managing environmental risks with a high spatiotemporal resolution and revealing environmental inequality across a diversity of population groups.

References

Barwick P J, Li S, Rao D, Zahur N B (2018). The healthcare cost of air pollution: Evidence from the world's largest payment network. Working Paper 24688. National Bureau of Economic Research

Baylis P (2020). Temperature and temperament: Evidence from Twitter. Journal of Public Economics, 184: 104161

Caplin A, Ghandehari M, Lim C, Glimcher P, Thurston G (2019). Advancing environmental exposure assessment science to benefit society. Nature Communications, 10(1): 1236

Casas L, Simons K, Nawrot T S, Brasseur O, Declerck P, Buyl R, Coomans D, Nemery B, van Nieuwenhuyse A (2016). Respiratory medication sales and urban air pollution in Brussels (2005 to 2011). Environment International, 94: 576–582

Chaix B (2018). Mobile sensing in environmental health and neighborhood research. Annual Review of Public Health, 39(1): 367–384

Colmer J, Hardman I, Shimshack J, Voorheis J (2020). Disparities in $PM_{2.5}$ air pollution in the United States. Science, 369(6503): 575–578

Deville Cavellin L, Weichenthal S, Tack R, Ragettli M S, Smargiassi A, Hatzopoulou M (2016). Investigating the use of portable air pollution sensors to capture the spatial variability of traffic-related air pollution. Environmental Science & Technology, 50(1): 313–320

Deville P, Linard C, Martin S, Gilbert M, Stevens F R, Gaughan A E, Blondel V D, Tatem A J (2014). Dynamic population mapping using mobile phone data. Proceedings of the National Academy of Sciences of the United States of America, 111(45): 15888–15893

Dewulf B, Neutens T, Lefebvre W, Seynaeve G, Vanpoucke C, Beckx C, van de Weghe N (2016). Dynamic assessment of exposure to air pollution using mobile phone data. International Journal of Health Geographics, 15(1): 14

Dons E, Laeremans M, Orjuela J P, Avila-Palencia I, Carrasco-Turigas G, Cole-Hunter T, Anaya-Boig E, Standaert A, de Boever P, Nawrot T, Götschi T, de Nazelle A, Nieuwenhuijsen M, Int Panis L (2017). Wearable sensors for personal monitoring and estimation of inhaled traffic-related air pollution: Evaluation of methods. Environmental Science & Technology, 51(3): 1859–1867

Fan Y, Palacios J, Arcaya M, Luo R, Zheng S (2021). Health perception and commuting choice: A survey experiment measuring behavioral trade-offs between physical activity benefits and pollution exposure risks. Environmental Research Letters, 16(5): 054026

Ghermandi A, Sinclair M (2019). Passive crowdsourcing of social media in environmental research: A systematic map. Global Environmental Change, 55: 36–47

Guo Y, Li Y (2018). Online amplification of air pollution risk perception: The moderating role of affect in information. Information Communication and Society, 21(1): 80–93

Hanigan I C, Rolfe M I, Knibbs L D, Salimi F, Cowie C T, Heyworth J, Marks G B, Guo Y, Cope M, Bauman A, Jalaludin B, Morgan G G (2019). All-cause mortality and long-term exposure to low level air pollution in the “45 and up study” cohort, Sydney, Australia, 2006–2015. Environment International, 126: 762–770

Hankey S, Lindsey G, Marshall J D (2017). Population-level exposure to particulate air pollution during active travel: Planning for low-exposure, health-promoting cities. Environmental Health Perspectives, 125(4): 527–534

Hong W, Wei Y, Wang S (2021). Left behind in perception of air pollution? A hidden form of spatial injustice in China. Environment and Planning C: Politics and Space, in press, doi:10.1177/23996544211036145

Liu M, Huang Y, Ma Z, Jin Z, Liu X, Wang H, Liu Y, Wang J, Jantunen M, Bi J, Kinney P L (2017). Spatial and temporal trends in the mortality burden of air pollution in China: 2004–2012. Environment International, 98: 75–81

Nyhan M, Grauwin S, Britter R, Misstear B, McNabola A, Laden F, Barrett S R H, Ratti C (2016). “Exposure Track”—The impact of mobile-device-based mobility patterns on quantifying population exposure to air pollution. Environmental Science & Technology, 50(17): 9671–9681

Ou Y, West J J, Smith S J, Nolte C G, Loughlin D H (2020). Air pollution control strategies directly limiting national health damages in the US. Nature Communications, 11(1): 957

Piel F B, Fecht D, Hodgson S, Blangiardo M, Toledano M, Hansell A L, Elliott P (2020). Small-area methods for investigation of environment and health. International Journal of Epidemiology, 49(2): 686–699

Qin J, Han S (2009). Neurocognitive mechanisms underlying identification of environmental risks. Neuropsychologia, 47(2): 397–405

Su J G, Barrett M A, Henderson K, Humblet O, Smith T, Sublett J W, Nesbitt L, Hogg C, van Sickle D, Sublett J L (2017). Feasibility of deploying inhaler sensors to identify the impacts of environmental triggers and built environment factors on asthma short-acting bronchodilator use. Environmental Health Perspectives, 125(2): 254–261

Sun C, Kahn M E, Zheng S (2017). Self-protection investment exacerbates air pollution exposure inequality in urban China. Ecological Economics, 131: 468–474

Tainio M, Jovanovic Andersen Z, Nieuwenhuijsen M J, Hu L, de Nazelle A, An R, Garcia L M T, Goenka S, Zapata-Diomedi B, Bull F, de Sá T H (2021). Air pollution, physical activity and health: A mapping review of the evidence. Environment International, 147: 105954

Tang R, Tian L, Thach T Q, Tsui T H, Brauer M, Lee M, Allen R, Yuchi W, Lai P C, Wong P, Barratt B (2018). Integrating travel behavior with land use regression to estimate dynamic air pollution exposure in Hong Kong. Environment International, 113: 100–108

Tsou M M, Lung S C C, Shen Y S, Liu C H, Hsieh Y H, Chen N, Hwang J S (2021). A community-based study on associations between $PM_{2.5}$ and PM_1 exposure and heart rate variability using wearable low-cost sensing devices. Environmental Pollution, 277: 116761

Vanroy C, Vanlandewijck Y, Cras P, Feys H, Truijen S, Michielsen M, Vissers D (2014). Is a coded physical activity diary valid for assessing physical activity level and energy expenditure in stroke patients? PLoS One, 9(6): e98735

Yang J, Qu S, Liu M, Liu X, Gao Q, He W, Ji J S, Bi J (2021). Gray cityscape caused by particulate matter pollution hampers human stress recovery. Journal of Cleaner Production, 279: 123215

Yue H, He C, Huang Q, Yin D, Bryan B A (2020). Stronger policy required to substantially reduce deaths from $PM_{2.5}$ pollution in China. Nature Communications, 11(1): 1462

Zheng S, Wang J, Sun C, Zhang X, Kahn M E (2019). Air pollution lowers Chinese urbanites’ expressed happiness on social media. Nature Human Behaviour, 3(3): 237–243