Environment International 104 (2017) 155–162

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

New frontiers for environmental epidemiology in a changing world

Cathryn Tonne[a,b,c,*], Xavier Basagaña[a,b,c], Basile Chaix[d,e], Maud Huynen[f], Perry Hystad[g], Tim S. Nawrot[h,i], Remy Slama[j], Roel Vermeulen[k,l], Jennifer Weuve[m], Mark Nieuwenhuijsen[a,b,c]

[a] Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain
[b] Universitat Pompeu Fabra (UPF), Barcelona, Spain
[c] CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
[d] Sorbonne Universités, UPMC Université, Nemesis research team, Paris 06, France
[e] INSERM, UMR_S 1136, Nemesis research team, France
[f] ICIS, Maastricht University, Maastricht, Netherlands
[g] College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, USA
[h] Centre for Environmental Sciences, Hasselt University, Hasselt, Belgium
[i] Centre for Environment & Health, Leuven University, Leuven, Belgium
[j] Team of Environmental Epidemiology applied to Reproduction and Respiratory Health, Inserm, CNRS, Univ. Grenoble-Alpes, IAB Joint Research Center, Grenoble, France
[k] Institute for Risk Assessment Sciences, Utrecht University, Utrecht, Netherlands
[l] Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, UK
[m] Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA

A R T I C L E I N F O

Keywords:
Environment
Epidemiology
Demographics
Technology
OMICs
Sensors

A B S T R A C T

Background: In the next 25 years, transformative changes, in particular the rapid pace of technological development and data availability, will require environmental epidemiologists to prioritize what should (rather than could) be done to most effectively improve population health.

Objectives: In this essay, we map out key driving forces that will shape environmental epidemiology in the next 25 years. We also identify how the field should adapt to best take advantage of coming opportunities and prepare for challenges.

Discussion: Future environmental epidemiologists will face a world shaped by longer lifespans but also larger burdens of chronic health conditions; shifting populations by region and into urban areas; and global environmental change. Rapidly evolving technologies, particularly in sensors and OMICs, will present opportunities for the field. How should it respond? We argue, the field best adapts to a changing world by focusing on healthy aging; evidence gaps, especially in susceptible populations and low-income countries; and by developing approaches to better handle complexity and more formalized analysis.

Conclusions: Environmental epidemiology informing disease prevention will continue to be valuable. However, the field must adapt to remain relevant. In particular, the field must ensure that public health importance drives research questions, while seizing the opportunities presented by new technologies. Environmental epidemiologists of the future will require different, refined skills to work effectively across disciplines, ask the right questions, and implement appropriate study designs in a data-rich world.

1. Introduction

Pekkanen and Pearce described the main challenges in environmental epidemiology in 2001 (Pekkanen and Pearce, 2001). These included complex mixtures of a large number of correlated exposures, small effect sizes which can lead to inconclusive studies in the context of residual confounding, and the need for new methods and interdisciplinarity to study links between global environmental change and health. They also warned that new technologies, rather than public health importance, might drive research questions (Pekkanen and Pearce, 2001). Although notable advances have since been made in statistical tools to derive useful information on mixtures of correlated exposures, (Agier et al., 2016; Bobb et al., 2015; Chadeau-Hyam et al., 2013) these, together with the precise measurement of exposures in space and time, remain some of the key challenges in the field today. The rapid pace of technological development and availability of data have made the need more acute to prioritize what should be done for maximum public health benefit over what could be done.

* Corresponding author at: ISGlobal, Doctor Aiguader, 88, 08003 Barcelona, Spain.
E-mail address: Cathryn.tonne@isglobal.org (C. Tonne).

http://dx.doi.org/10.1016/j.envint.2017.04.003
Received 21 January 2017; Received in revised form 7 April 2017; Accepted 10 April 2017
Available online 26 April 2017

In Part I of this essay, we map out key driving forces that will shape environmental epidemiology in the next 25 years (a time span roughly overlapping with most of the authors´ careers). In Part II, we suggest how the field can leverage these forces to improve population environmental health. Specifically, we suggest ways in which new paradigms, measurement strategies, and analytical approaches can be adapted. Among possible future scenarios, we highlight those which we believe are preferable for achieving public health goals.

1.1. Part I. Forces that will shape the field

1.1.1. Demographics and urbanization

Within the broader context of the epidemiologic transition and economic development, the changing age and geographic distribution of the global population will continue transforming environmental health research priorities. Lifespans have lengthened across the globe, including important gains in low-income countries where life expectancy has improved from 53 years in 1990 to 62 years in 2012 (WHO, 2014). Adults aged 60 and older comprised 8% of the global population in 1940, which grew to 12% by 2013 and is projected to be 21% in 2050 (UN Department of Economic and Social Affairs, 2013). This monumental shift in the global age distribution is due in part to public health interventions, but also presents new health challenges.

There will also be notable shifts in the geographic distribution of the global population. Most of the population growth between now and 2050 is projected to occur in just nine countries with high fertility or already large populations; more than half of the expected growth will be in Africa alone (UN Department of Economic and Social Affairs, 2015). The proportion of the population living in Africa will increase from 16% to 25% between now and 2050, while the proportion in Europe will shrink from 10% to 7% (UN Department of Economic and Social Affairs, 2015). Population growth is projected to remain especially high in the 48 least developed countries, adding to challenges in meeting sustainable development goals (UN Department of Economic and Social Affairs, 2015). There is currently very little data on environmental exposures and their health effects in many countries with the largest population growth. Effects of environmental exposures are likely to differ from those observed in high-income countries, where most environmental health research has been conducted, due to differences in infectious disease burden, access to health services, and material deprivation.

Large-scale migration will add further complexity for environmental epidemiology, presenting challenges for follow-up of study participants and environmental exposure assessment. Individuals living in areas with similar levels of environmental exposures may have highly variable cumulative exposure based on their migration history. For migrants from poorly to better regulated societies, adult health may be influenced by high levels of environmental exposures in early life, exposures which may be particularly difficult to reconstruct. Migration may also present opportunities for using natural experiments to understand how environment shapes health.

The global population continues to shift from rural to urban areas. In 2014, 54% of the population resided in urban areas; this is projected to be 66% by 2050 (UN Department of Economic and Social Affairs, 2014). Nearly 90% of the projected increase in the world's urban population will be concentrated in Asia and Africa, with India, China, and Nigeria accounting for a large share of this growth. Urbanization profoundly shapes (both positively and negatively) environmental exposures (e.g. air pollution, noise, green space) and behaviors (e.g. physical activity, food consumption) and thereby disease risks (Nieuwenhuijsen, 2016).

1.1.2. Global environmental change

Climate change and emerging environmental risks will define much of the future context for environmental epidemiology. Climate change has been identified as "the biggest global health threat of the 21st century" (Costello et al., 2009). Mean surface temperature is expected to increase by 0.3 to 4.8 °C by 2100, (IPCC, 2013) leading to direct impacts on health from heat stress and flooding, as well as indirect health impacts mediated through infectious diseases, air quality, and food security (Anstey, 2013; Costello et al., 2009; IPCC, 2014; McMichael, 2012; Patz et al., 2016).

Recent reports have elaborated the multiple potential health impacts of climate change (Whitmee et al., 2015). There is strong evidence that heat-related mortality is rising as a result of climate change; the Intergovernmental Panel on Climate Change anticipates an increase in both the frequency and intensity of heat waves under all climate scenarios (IPCC, 2014). The combined effect of global warming and demographic change will expose an increasing number of vulnerable older adults to heat stress (Watts et al., 2015). Health impacts of extreme weather events such as storms and floods are likely to increase this century if no adaptation measures are taken (IPCC, 2014). It is also anticipated that climate change increases the risk of intense droughts in some regions, (IPCC, 2013) affecting agricultural output and, subsequently, increasing food insecurity and malnutrition. Climate change has been identified as one of the greatest challenges for food security (High Level Panel of Experts on Food Security and Nutrition, 2012). Droughts also elevate risks of water-related disease (e.g. *E coli*, cholera), vector borne disease (e.g. dengue, West Nile Virus), airborne and dust-related disease (e.g. coccidioidomycosis) and mental illness (Stanke et al., 2013). Climate (change) may be an important factor in the dynamics of vector borne disease transmission, including malaria, dengue, and Lyme's disease (Chaves and Koenraadt, 2010; IPCC, 2014). Alongside shifts in land use, climate change appears to be altering the geographic range of vectors that transmit pathogens (e.g. *Aedes albopictus*, *Aedes albopitcus, Ixodes scapularis*) to humans (Murray, 2013; Ogden and Radojevic, 2014; Proestos et al., 2015).

Chemical exposures will remain an important environmental health concern. Chemical production in 2000 was 1000 times higher than in 1930 (UNEP, 2013). Although chemical production is not a direct measure of population exposure, it is likely that the number of chemicals to which one is exposed will continue to increase in coming decades. Of particular concern are those with short half-lives in the body, now preferred to those with long half-lives for health and environmental reasons, but which contribute to exposure misclassification in traditional studies that rely on spot biomarkers (Perrier et al., 2016).

1.1.3. Technology

Technology with applications to environmental exposure and health outcome assessment is evolving rapidly. Technology will generate new opportunities, particularly in regards to population datasets, e- and m (obile)-health, personal and remote sensor technology, and OMICs data. Below we highlight technologies that lie on the horizon and how they could be applied to environmental epidemiology.

Expanding data availability will allow prediction of diverse population exposures and create new opportunities for exploring novel exposures that have been previously difficult to quantify. Importantly, geo-referenced data are becoming more widely available in low- and middle-income countries, reducing barriers for conducting environmental epidemiology in these countries. Such data include those collected through remote sensing, sensor networks, smartphones, as well as the "internet of things" (i.e., everyday objects with network connectivity). Remote sensing has been used to estimate environmental exposures including air pollution, (Geddes et al., 2016) green space, (Dadvand et al., 2015) and temperature (Dadvand et al., 2014). Opportunities for satellite-based exposure assessment will continue to expand with increasing number of satellites and improved resolution of detection. Quantifying neighborhood attributes will be enhanced by applying developments in image processing to resources such as Google Street View and to ecological momentary assessment based on individuals taking a photograph of their immediate environment from their mobile phone. (Advances in image processing will also improve measurement of other exposures, such as diet or drug or cosmetics use, for which study participants can take pictures of what they eat or use, or scan bar codes). A new exposure pathway − visual exposure − - will be easily investigated using miniaturized cameras or virtual reality to understand how people internalize and interact with their environment. For example, such technology will allow advances beyond simple proximity to green space to determine whether individuals are visually exposed to green space and which activities they engage in using that green space. Social media data will increasingly play a role in assessing behaviors, exposures and outcomes. Such approaches have already been used to identify symptoms, behavioral risk factors, and population mobility patterns (Lee et al., 2016; Paul and Dredze, 2011).

The rise of e/m-health and new personal sensor technologies is opening tremendous opportunity for environmental epidemiology. Recruiting participants, collecting survey data, assessing outcomes, and incorporating medical health records using mobile platforms (e.g. smart phones) is becoming easier and will be the core mechanism used to recruit cohorts (e.g. the NIH precision medicine cohort of 1 million individuals) (NIH, 2016). The use of e/m-health is further propelled forward by the quantified self movement (self tracking with technology) (Swan, 2013). The growing acceptability of existing smart phones and sensors foreshadow the types and size of data that may be available when low-cost sensors for environmental exposures improve in terms of data quality and continue to decrease in size and cost. Inexpensive wearable devices can now register numerous biological parameters in real time, including participants' physical activity, sleep, heart rate, and, to a limited extent, blood pressure, with little participant burden. These types of monitors will allow measurement of exposures and health outcomes on larger scales for research purposes. As the price of sensors drops, they are also likely to be widely used by individuals outside of research contexts. Smart phone manufacturers will also recognize the marketing value of differentiating their devices by adding sensor packages.

In 25 years, new sensors will have completely revolutionized environmental exposure and health outcome assessment. These sensors will measure more things and give constant feedback to individuals, with important implications for observational study design. There will also be easy access to new sensors through widespread 3D printing capabilities, revolutionizing the concept of citizen science. The availability and connectedness of disparate data-sources, the “internet of things”, will be a driving force shaping exposure assessment. Homes, street furniture, and cars will be increasingly equipped with smart sensors to monitor environmental conditions (e.g. air pollutants, asthma triggers, noise, etc.); smart cities will provide continuous data streams, and crowd sourcing of personal sensor measurements will be integrated within these environment measures. Environmental measurements will be an established component of self-tracking, along with health behaviors and a growing scope of biological measures (e.g. wearable patches that transmit continuous heart rates measurements, blood chemistry, stress levels, etc.). Wider access to and use of smart devices could enhance the frequency of participation and reduce selection bias where participation is linked to exposure (Weisskopf et al., 2015). However, important challenges around data ownership and accessibility will certainly arise; legal and ethical norms will struggle to keep pace with what is technologically possible. There is an obvious need for closer engagement between environmental epidemiologists and the ethics community to provide practical, timely guidance on how to take advantage of the opportunities for improving population environmental health in an ethical way.

The paradigm of individuals having information drawn from them may change to individuals being suppliers of information, choosing how and with whom to share their data. Potential incentives to individuals to “donate” their data to a central repository for research purposes may include discounts on cell phone bills, tax write-offs, etc. - processes that could lead to selection bias. Issues related to the validation and quantification of exposure misclassification induced by the use of heterogeneous measurement devices should not be underestimated and are expected to be even more complex than when a single measurement device is used for all participants.

Similarly, important advances have been made recently in OMICS technologies (i.e. any group of measurements covering the totality (or large proportion) of a dimension of an environmental, behavioral, social, or biological variable, (e.g., genomics, epigenomics, transcriptomics, proteomics, adductomics, metabolomics, microbiomics, exposomics) that could be used to supplement the assessment of internal exposures by more traditional means (Pedersen and Nieuwenhuijsen, 2015). These technologies are rapidly decreasing in cost and can increasingly be applied to larger populations for multiple purposes. Current applications of OMICs technologies provide an indication of what will be possible as the technology improves and costs decrease further. An integrated personal OMICs profile combining genomic, transcriptomic, proteomic, metabolomic and autoantibody profiles has been constructed for a single individual over 14 months, revealing dynamic phenotypes as well as internal exposures (Chen et al., 2012). Although, it is not currently possible to measure personal OMICs profiles in such a dense format on large numbers of participants, studies have started to measure repeated OMICs profiles in populations with a few thousand participants (Vrijheid et al., 2014). OMICs profiles have been used to measure exposure directly or by measuring the imprint that environmental factors leave in the biological system. For example, transcriptomics and methylation patterns have recently been used to identify participants´ smoking history (Beane et al., 2007; Guida et al., 2015) and exposure to smoky vs. smokeless coal (Wang et al., 2015b). Further rapid advances are expected through current exposome projects (e.g. HELIX, EXPOsOMICs, HEALS, and CHEAR).(CHEAR, n.d.; HEALS, n.d.; Vineis et al., 2016; Vrijheid et al., 2014).

Projections based on current advances in OMICs technologies, suggest that by 2035–45, we should be able to measure an individual's full OMICs profile (Fig. 1). Repeated OMICs profiles would open enormous possibilities for life-course studies, natural experiments, and assessing the long-term impacts of interventions. It would also allow for systematic screening of known, suspected and unknown agents/phenomes in a holistic way. However, interpretation will become the major challenge as our biological knowledge may not keep up with the technology. Similarly, statistical and bio-informatics capabilities could become rate limiting and will require rapid further development.

1.1.4. Data availability

Data creation is already exceeding worldwide storage capacity; (Harcourt, 2014) this may become a problem in epidemiologic studies. Table 1 illustrates the size of different types of data currently collected in epidemiologic studies. Genomics is a familiar example where data size is already an issue. The addition of other OMICs escalates the problem (e.g. the human microbiome would contain the genomic information of 100 trillion cells). New technologies (e.g. sensors, medical image and video) are also able to provide huge amounts of data for a single participant. This movement towards Big Data studies will make parallel processing on computer clusters or a cloud, and the use of Big Data platforms more common. In many instances, raw data will be discarded owing to storage problems and only relevant summaries will be stored. Big data created for purposes other than research are likely to present challenges in terms of data quality and representativeness of the wider population.

1.1.5. Study design, models of research

We will increasingly move towards studies with very large sample sizes with individual exposure information, even including the entire population (N = all) as is currently possible with census cohorts. One may think that statistics are no longer needed if the entire population is observed, but this is probably the case only when computing simple

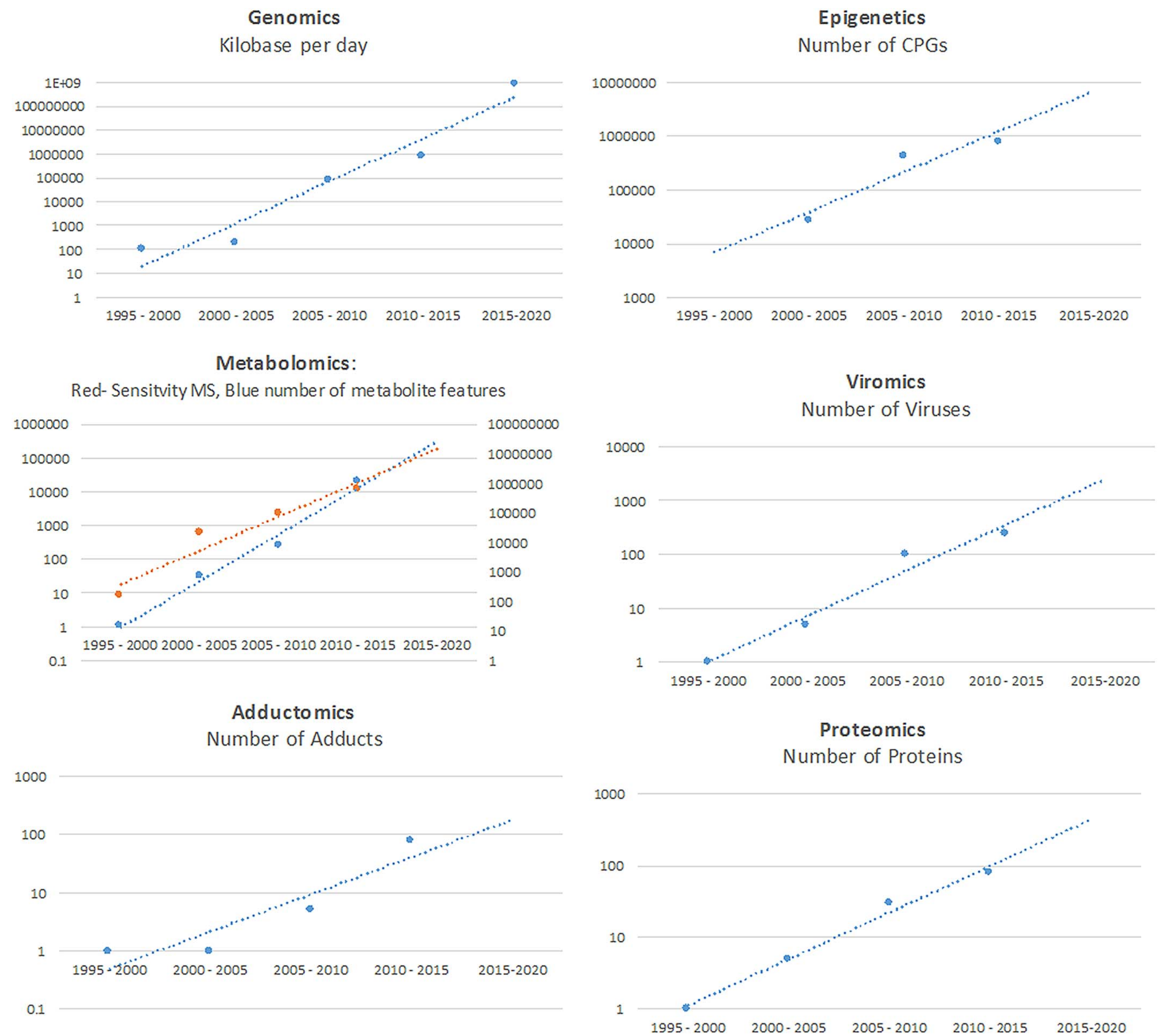

Fig. 1. Predicted growth over time in internal exposure assessment capacity.

summary statistics (e.g. a mean). Environmental epidemiology often deals with complex causal questions, which require the observation of a set of confounders. Empirical estimations of those effects would require calculating averages over all possible combinations of confounders. With just 10 confounders, this becomes unfeasible with huge datasets, because of empty cells, so statistical models will continue to be needed (van der Laan and Starmans, 2014). In addition, if the entire population is observed, the traditional statistical paradigm based on the sampling of independent observations no longer holds. Informative dependencies exist in the population, and some authors suggest that the population's network structure should be taken into account when analyzing whole-population data (van der Laan and Starmans, 2014). More importantly, considering the expected explosion in OMICs and sensor data, future studies will tend to have a much larger number of variables than participants (so-called *high-dimension* data). Most of the statistical methods commonly used today such as maximum likelihood estimation do not provide consistent estimates in such settings (Fan et al., 2014). Large dimensionality brings spurious correlations, among other problems (Fan et al., 2014). Techniques for dimension reduction, variable selection, and sparse models that can work effectively in that setting will become crucial.

Other models of research production will come online. Citizen science, to date most widely used in fields of biology and ecology, is likely to play a greater role in environmental epidemiological research production and may offer benefits in terms of better knowledge, empowered communities, and improved health outcomes (Den Broeder et al., 2016). However, these benefits have yet to be fully evaluated and several practical challenges remain (Den Broeder et al., 2016). The current model of research production encourages publication of single analyses, often performed by a single research group, which can be dominated by subjective decisions during the analysis and favor extreme results. Moving towards a researcher-based crowd-sourced analysis model potentially offers many advantages, including greater transparency, opportunity to vet analytical approaches with peers before publication, and more nuanced, balanced results (Silberzahn and Uhlmann, 2015). The Many Lab, launched by the Centre for Open Science, is a current example of a web-platform for crowdsourced research production where researchers can join projects

Table 1
Examples of storage needs (per person).

Information	Size
Human genome	1 GB
+ structure determination of the proteins	Several PB[a]
Electronic health record	1 MB-5GB, expected to increase 50-fold from 2012 to 2020[b,c]
Heart rate monitor (per month)	9 GB[d]
Continuous video life-logger (per month)	58GB[e]
Accelerometer (8-h a day, per month)	1 GB[f]
Medical image	MB to GB, up to 1 TB. e.g. 64/128-slice CT scan, 3.0 T MRI and PET often exceeding 100 MB[b].

MB = 10^6 bytes; GB = 10^9 bytes; TB = 10^{12} bytes; PB = 10^{15} bytes.
[a] Al-Jarrah OY, Yoo PD, Muhaidat S, Karagiannidis GK, Taha K. 2015. Efficient machine learning for big data: a review. Big Data Research, 2(3), 87–93.
[b] Piai S, Claps, M. 2013. Bigger data for better healthcare. IDC Health Insights. Available: http://www.intel.com/content/www/us/en/healthcare-it/solutions/documents/bigger-data-better-healthcare-idc-insights-white-paper.html [accessed 16 February 2016]
[c] Radding A. 2008. Storage gets a dose of medical data. Storage Magazine. Available: http://searchstorage.techtarget.com/magazineContent/Storage-gets-a-dose-of-medical-data [accessed 16 Feb 2016]
[d] Swan M. 2013. The quantified self: Fundamental disruption in big data science and biological discovery. Big Data, 1(2), 85–99.
[e] Schneier, Bruce. Data and Goliath: The hidden battles to collect your data and control your world. WW Norton & Company, 2015.
[f] Baumann, L., Kesztyüs, T., & Blechschmidt, R. A. Standard compliant communication of motion data in a telemonitoring system. GMS Medizinische Informatik, Biometrie und Epidemiologie 2015, 11(1), 1–12.

(Center for Open Science, n.d.). The advantages of research involving large groups should be weighed against bureaucratic inefficiencies that hamper individual creativity and curiosity-driven fundamental research.

1.2. Part II. How environmental epidemiology should adapt to take advantage of coming opportunities and prepare for future challenges

1.2.1. Prioritize healthy aging

Environmental epidemiology should respond to demographic trends by increasing focus on healthy aging. By not succumbing to the diseases of infancy, childhood, and middle age, a growing number of people are affected by chronic health conditions of older age. Worldwide, all-cause dementia affected 36 million in 2010, a number anticipated to skyrocket to 115 million by 2050 (Prince et al., 2013). Even more common in older adulthood is physical disability, (WHO and NIH, 2011) which afflicts over 40% of older adults in the US, England and Europe (Avendano et al., 2009). In 188 countries evaluated by the Global Burden of Disease Study (GBD), life expectancy increased between 1990 and 2013, but so did the years lived with disability (Global Burden of Disease Collaborators, 2015). Although the burden of these conditions has been a paradoxical luxury of living in a high-income country, low and middle-income countries will bear more of the brunt of the dementia burden, (WHO and NIH, 2011) and the same can be expected of physical disability.

Aging is universal and unavoidable, but the aging process does not occur uniformly. It is a complex phenotype responsive to a plethora of drivers. Genetic, behavioral and environmental factors interact with each other to define an individual's aging trajectory. If environmental exposures influence the development of dementia, physical disability or even the many other common conditions of older age, the potential benefit of shifting these epidemic curves downward is enormous due to the high prevalence of both the outcomes and many environmental exposures. Better understanding of the effects of environmental exposures over the life course on health in older ages is critical to effective prevention strategies, as disease onset may start in early life (Vrijheid et al., 2014).

1.2.2. Increase equity of information

Environmental epidemiology should prioritize improving the equitable distribution of data on environmental exposures and their health effects, and maximizing the societal benefit of such data. Global, systematic research initiatives such as the GBD have highlighted dramatic inequities across populations in the availability of exposure and epidemiological evidence. For example, early iterations of the GBD highlighted the relative lack of outdoor air pollution exposure information for most of the global population (Cohen et al., 2004). The exposure-response estimate from a single large US population was assumed to apply to the entire global population, even though that study covered a fraction of the global exposure range (Cohen et al., 2004). Such efforts have put into focus the need for better exposure and epidemiological information across the true global exposure range. Low-cost, scalable data collection methods, particularly through remote sensing and small sensors, offer an important opportunity to achieve better equity of information on environmental exposures, health outcomes, and epidemiological evidence, particularly for rural, low-income, or other populations currently living in "information deserts".

Systematic data registration will uncover gaps in available environmental exposure data and exposure-response estimates. Meta-research, or research on research, can push forward systematic mapping of gaps in the environmental epidemiology evidence base (Khoury et al., 2013). Decision analysis approaches will be useful in prioritizing particular populations for which additional research would be most scientifically valuable and policy relevant. Populations to target would be those outside of previously studied exposure ranges, unique combinations of exposures, or high likelihood for effect modification.

The full potential of the vast exposure data that will be collected in the future is best realized when those data are open (in a sufficiently anonymized form to protect privacy). Open exposure data will allow for use, reuse, and redistribution of data with proper attribution. Open data platforms that provide standards of data quality could be used as a global repository for crowd sourced exposure data. A useful example in this regard is Open Street Map, a collaborative project to create a free, editable, global geodatabase. Although currently limited to data from official, stationary government monitors, openaq.org is another example of a collaborative project to aggregate, standardize, and share real-time air quality data from around the world. Platforms that standardize and harmonize environmental data globally will play an increasingly important role in maximizing value from environmental exposure data for societal gain. Mechanisms to support (financially and otherwise) the curation of these databases should be further developed.

1.2.3. Reduce participant burden

Data collection methods should become increasingly passive to minimize the burden to participants. Already several countries allow linkage of administrative and registry data for research purposes, allowing for epidemiology on entire populations with minimal participant burden. Response rates in many population studies are falling compared to past decades; this has been cited as a possible factor leading to the cancellation of a major birth cohort study in the UK due to low recruitment (Pearson, 2015). By embedding data collection in routine medical care, longitudinal OMICs data could be collected with little additional burden. Potentially thousands of analytes could be measured on each blood or urine sample collected as part of clinical care; access to these samples for research would be an important way to enable lifecourse epidemiology.

1.2.4. Expand approaches to handle complexity

The traditional and currently dominant epistemological approach in epidemiology is reductionism: individual exposures are studied rather than the system of health determination as a whole or the causal architecture (Keyes and Galea, 2017). Study designs focus on precisely estimating individual causal-effect relationships rather than exploring system interactions. Moving towards system-based approaches would allow for possible interactions between different exposures, biological parameters and developments (including urbanization, climate change adaptation) (Huynen et al., 2013; Rydin et al., 2012; Whitmee et al., 2015). Single exposures and regression-based analysis methods should be replaced with approaches better equipped to deal with multiple exposures, as is envisaged in the exposome and consequential epidemiology approaches (Keyes and Galea, 2017; Wild, 2012).

Improving integrated risk assessments in environmental epidemiology should be a priority, particularly for the link between climate change and health. Traditional forms of risk assessment do not adequately assess systemic and/or longer-term global environmental health risks (Briggs, 2008). For example, most climate-related health impacts are mediated through complex ecological and social processes (Watts et al., 2015). As such, climate change affects health through extensive and complex linkages between multitudes of factors, which together determine the vulnerability-context of a specific population. Through these system interactions, climate change can often act as an important amplifier of existing health risks (Costello et al., 2009; Huynen et al., 2013; WHO, 2009). In line with such broader approaches, risk assessments should, at a minimum, explore the negative and positive impacts of climate-related mitigation and adaptation policies in other sectors (Milner et al., 2014; Sabel et al., 2016; WHO, 2009).

1.2.5. Formalize data analysis

Data analysis and consequently epidemiology, should go through a formalization process to become a less subjective process and to prevent publication of analyses lacking theoretical rigor (van der Laan, 2015). This includes explicitly stating the assumed causal structure (e.g. through causal directed acyclic graphs) and assumptions, which can be particularly challenging in the context of repeated measures of time-varying OMICs. Methods for evaluating a finding's sensitivity to different forms of bias and departures from assumptions are coming into wider use (Power et al., 2016). These methods will become a pivotal tool for overcoming a frequent obstacle in environmental epidemiology, which is that randomized controlled trials are infrequently possible or ethical. Appropriate methods for causal inference should, as standard procedure, include techniques to correct for non-participation, missing data, measurement error and multiple testing. Semi-parametric and non-parametric methods should be favored, as parametric models make unrealistic simplifying assumptions that are believed to produce invalid inferences (Rudin et al., 2014; van der Laan and Starmans, 2014). Other future challenges for statisticians and epidemiologists include being able to formulate testable hypotheses from poorly defined and ambiguous problems involving increasingly dense data from different fields (Wang et al., 2015a).

1.2.6. Adapt training of future environmental epidemiologists

To maintain relevance as a field, future environmental epidemiologists must be prepared for the coming changes, a topic recently discussed in the context of epidemiology more broadly (Brownson et al., 2016). Big Data will drive many changes in the way we work and analyze data. With Big Data, even simple models involve large amounts of computation and some become computationally intractable (Rudin et al., 2014). Future epidemiologists will require training in new techniques to scale existing methods for Big Data, and new statistical methods to overcome limitations of existing ones. New methods for knowledge discovery will be developed in other fields (e.g. machine learning, applied mathematics, statistics) which are likely to crossover into epidemiology. Future data analysts will need to be familiar with those fields and with large-scale computing (e.g. parallel processing). As in the past, trainees will require strong backgrounds in epidemiology methods and statistics; however these areas will also change to adopt different methods and even priorities.

Epidemiologist of the future will be part of transdisciplinary teams, working alongside social scientists, biologists, toxicologists, pathologists, computer scientists, urban planners, and specialists in knowledge translation. Training should include more focus on implementation, evaluation, and intervention. In parallel, there is an argument to shift the focus of some environmental epidemiology from ever more precise estimation of individual risk factors to what matters most to population health (Keyes and Galea, 2015).

2. Conclusions

A large amount of disease etiology remains unexplained and is likely to be due to the environment (Norman et al., 2013; Willett, 2002). Many transformative changes are expected in the next 25 years, and environmental epidemiologists will continue to have a role to play in disease prevention.

However, the field must adapt in order to remain relevant. New technologies are inevitable and many present important opportunities for environmental epidemiology. The previously cited risk of new technologies driving research questions rather than public health importance (Pekkanen and Pearce, 2001) will become increasingly relevant. The challenge will be how best to harness these technologies for environmental health goals.

With increasing data availability there will be greater need for robust and new study designs to identify causal associations in a data-rich world full of false positives. The distinction between population and individual exposure and outcome assessment will erode due to the magnitude of individual-level data available through sensors interfaced with smart technologies and biotechnologies. The environmental epidemiologist of the future will require different and refined skills to work effectively across disciplines, ask the right questions, implement the most appropriate designs and analyses methods, to identify causal relationships, ignite and engage interest of the public and policy makers, and design effective interventions. If successful, environmental epidemiology will continue to develop novel preventive strategies that save health care costs and add years and quality to life at the population level.

Sources of financial support

This manuscript is based on a symposium held at the Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain on 27 November 2015.

Funding for the symposium was provided by CREAL.

Conflicts of interest

J.W. serves as a consultant to the Alzheimer's Association, Biogen, Inc. and the AlzRisk Project (http://www.alzrisk.org), which receives funding from Fidelity Biosciences and from the Alzheimer's Drug Discovery Foundation. The work entailed interpretation of published dementia epidemiology studies. All authors declare they have no actual or potential competing financial interests.

References

Agier, L., Portengen, L., Chadeau-Hyam, M., Basagana, X., Giorgis-Allemand, L., Siroux, V., et al., 2016. A systematic comparison of linear regression-based statistical methods to assess exposome-health associations. Environ. Health Perspect. http://dx.doi.org/10.1289/EHP172.

Anstey, M.H.R., 2013. Climate change and health – what's the problem? Glob. Health 9, 1–5.

Avendano, M., Glymour, M.M., Banks, J., Mackenbach, J.P., 2009. Health disadvantage in US adults aged 50 to 74 years: a comparison of the health of rich and poor Americans with that of Europeans. Am. J. Public Health 99, 540–548.

Beane, J., Sebastiani, P., Liu, G., Brody, J.S., Lenburg, M.E., Spira, A., 2007. Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene

expression. Genome Biol. 8, R201.
Bobb, J.F., Valeri, L., Claus Henn, B., Christiani, D.C., Wright, R.O., Mazumdar, M., et al., 2015. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostatistics 16, 493–508.
Briggs, D.J., 2008. A framework for integrated environmental health impact assessment of systemic risks. Environ. Health 7, 61.
Brownson, R.C., Samet, J.M., Bensyl, D.M., 2016. Applied epidemiology and public health: are we training the future generations appropriately? Ann. Epidemiol. http://dx.doi.org/10.1016/j.annepidem.2016.12.002.
Center for Open Science The many lab. https://osf.io/89vqh/ (accessed 15 August 2016).
Chadeau-Hyam, M., Campanella, G., Jombart, T., Bottolo, L., Portengen, L., Vineis, P., et al., 2013. Deciphering the complex: methodological overview of statistical models to derive OMICS-based biomarkers. Environ. Mol. Mutagen. 54, 542–557.
Chaves, L.F., Koenraadt, C.J.M., 2010. Climate change and highland malaria: fresh air for a hot debate. Q. Rev. Biol. 85, 27–55.
CHEAR Children's health exposure analysis resource. https://www.niehs.nih.gov/research/supported/exposure/chear/ (Accessed 15 Aug 2016).
Chen, R., Mias, G.I., Li-Pook-Than, J., Jiang, L., Lam, H.Y.K., Chen, R., et al., 2012. Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell 148, 1293–1307.
Cohen, A.J., Anderson, H.R., Ostro, B., Pandey, K.D., Krzyzanowski, M., Kunzli, N., et al., 2004. Urban air pollution. In: Comparative Quantification of Health Risks. Vol. 2 WHO.
Costello, A., Abbas, M., Allen, A., Ball, S., Bell, S., Bellamy, R., et al., 2009. Managing the health effects of climate change: Lancet and University College London Institute for Global Health Commission. Lancet 373, 1693–1733.
Dadvand, P., Ostro, B., Figueras, F., Foraster, M., Basagaña, X., Valentin, A., et al., 2014. Residential proximity to major roads and term low birth weight: the roles of air pollution, heat, noise, and road-adjacent trees. Epidemiology 25, 518–525.
Dadvand, P., Nieuwenhuijsen, M.J., Esnaola, M., Forns, J., Basagaña, X., Alvarez-Pedrerol, M., et al., 2015. Green spaces and cognitive development in primary schoolchildren. Proc. Natl. Acad. Sci. 201503402.
Den Broeder, L., Deville, J., Van Oers, H., Schuit, A.J., Wagemakers, A., 2016. Citizen Science for public health. Health Promot. Int. 1–10. http://dx.doi.org/10.1093/heapro/daw086.
Fan, J., Han, F., Liu, H., 2014. Challenges of big data analysis. Natl. Sci. Rev. 1, 293–324.
Geddes, J.A., Martin, R.V., Boys, B.L., van Donkelaar, A., 2016. Long-term trends worldwide in ambient NO2 concentrations inferred from satellite observations. Environ. Health Perspect. 124 (3), 281–289.
Global Burden of Disease Collaborators, 2015. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study. Lancet 386, 743–800.
Guida, F., Sandanger, T.M., Castagne, R., Campanella, G., Polidoro, S., Palli, D., et al., 2015. Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. Hum. Mol. Genet. 24, 2349–2359.
Harcourt, B., 2014. Governing, exchanging, securing: big data and the production of digital knowledge. In: Columbia Public Law Research Paper, pp. 14–390.
HEALS Health and environment-wide associations based on large population surveys. http://www.heals-eu.eu (Accessed 15 Aug 2016).
High Level Panel of Experts on Food Security and Nutrition, 2012. Food security and climate change. Available:. http://www.fao.org/fileadmin/user_upload/hlpe/hlpe_documents/HLPE_Reports/HLPE-Report-3-Food_security_and_climate_change-June_2012.pdf (accessed 15 July 2016).
Huynen, M.M.T.E., Martens, P., Akin, S., 2013. Climate change: an amplifier of existing health risks in developing countries. Environ. Dev. Sustain. 15, 1425.
IPCC, 2013. Climate change 2013: the physical science basis: summary for policymakers, intergovernmental panel on climate change. In: Working Group I Contribution to the IPCC Fifth Assessment Report, .
IPCC, 2014. Climate change 2014: impacts, adaptation and vulnerability. In: Working Group II Contribution to the IPCC Fifth Assessment Report, .
Keyes, K., Galea, S., 2015. What matters most: quantifying an epidemiology of consequence. Ann. Epidemiol. 25, 305–311.
Keyes, K.M., Galea, S., 2017. Commentary: the limits of risk factors revisited: is it time for a causal architecture approach? Epidemiology 28 (1).
Khoury, M.J., Lam, T.K., Ioannidis, J.P.A., Hartge, P., Spitz, M.R., Buring, J.E., et al., 2013. Transforming epidemiology for 21st century medicine and public health. Cancer Epidemiol. Biomark. Prev. 22, 508–517.
van der Laan, M.J., 2015. Statistics as a science, not an art: the way to survive in data science. Available. http://magazine.amstat.org/blog/2015/02/01/statscience_feb2015/ (accessed Aug 15 2016).
van der Laan, M.J., Starmans, R.J.C.M., 2014. Entering the era of data science: targeted learning and the integration of statistics and computational data analysis. Adv. Stat. 2014, 502678. http://dx.doi.org/10.1155/2014/502678.
Lee, J.H., Gao, S., Goulias, K.G., 2016. Can Twitter data be used to validate travel demand models? Available:. http://www.geog.ucsb.edu/~sgao/papers/2016_TRB_TweetOD.pdf (Accessed 15 August 2016).
McMichael, A.J., 2012. Insights from past millennia into climatic impacts on human health and survival. Proc. Natl. Acad. Sci. 109 (13).
Milner, J., Shrubsole, C., Das, P., Jones, B., Ridley, I., Chalabi, Z., et al., 2014. Home energy efficiency and radon related risk of lung cancer: modeling study. BMJ. http://dx.doi.org/10.1136/bmj.f7493.
Murray, A.N.E., 2013. Epidemiology of dengue: past, present and future prospects. Clin. Epidemiol. 5, 299–309.
Nieuwenhuijsen, M.J., 2016. Urban and transport planning, environmental exposures and health-new concepts, methods and tools to improve health in cities. Environ. Health 15 (S1), S38.
NIH, 2016. Precision medicine initiative cohort program. Available:. https://www.nih.gov/precision-medicine-initiative-cohort-program/scale-scope (Accessed 15 Aug 2016).
Norman, R.E., Carpenter, D.O., Scott, J., Brune, M.N., Sly, P.D., 2013. Environmental exposures: an underrecognized contribution to noncommunicable diseases. Rev. Environ. Health 28, 59–65.
Ogden, N.H., Radojevic, M., 2014. Estimated effects of projected climate change on the basic reproductive number of the Lyme disease vector Ixodes scapularis. Environ. Health Perpect. 122, 631–638.
Patz, J., Frumkin, H., Holloway, T., Vimont, D., Haines, A., 2016. Climate change challenges and opportunities for global health. JAMA 53726, 1565–1580.
Paul, M.J., Dredze, M., 2011. You are what you tweet: analyzing Twitter for public health. Available. https://www.cs.jhu.edu/~mdredze/publications/twitter_health_icwsm_11.pdf (accessed Aug 15 2016).
Pearson, H., 2015. Massive UK baby study cancelled. Nature 526, 620–621. http://dx.doi.org/10.1038/526620a.
Pedersen, M., Nieuwenhuijsen, M.J., 2015. Biological monitoring. In: Nieuwenhuijsen, M.J. (Ed.), In Exposure Assessment in Environmental Epidemiology. Oxford Univ Press.
Pekkanen, J., Pearce, N., 2001. Environmental epidemiology: challenges and opportunities. Environ. Health Perspect. 109, 1–5.
Perrier, F., Giorgis-Allemand, L., Slama, R., Philippat, C., 2016. Within-subject pooling of biological samples to reduce exposure misclassification in biomarker-based studies. Epidemiology 27, 378–388.
Power, M.C., Adar, S.D., Yanosky, J.D., Weuve, J., 2016. Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: a systematic review of epidemiologic research. Neurotoxicology 56, 235–253. http://dx.doi.org/10.1016/j.neuro.2016.06.004.
Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., Ferri, C.P., 2013. The global prevalence of dementia: a systematic review and metaanalysis. Alzheimers Dement. 9, 63–75 (e2).
Proestos, Y., Christophides, G.K., Ergu, K., Tanarhte, M., Waldock, J., Lelieveld, J., 2015. Present and future projections of habitat suitability of the Asian tiger mosquito, a vector of viral pathogens, from global climate simulation. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 370 (1665) (5).
Rudin, C., Dunson, D., Irizarry, R., Ji, H., 2014. Discovery with data: leveraging statistics with computer science to transform science and society. American Statistical Association, white paper. Available. https://www.amstat.org/policy/pdfs/BigDataStatisticsJune2014.pdf (accessed July 15 2016).
Rydin, Y., Bleahu, A., Davies, M., Dávila, J.D., Friel, S., De Grandis, G., et al., 2012. Shaping cities for health: complexity and the planning of urban environments in the 21st century. Lancet 379, 2079–2108.
Sabel, C.E., Hiscock, R., Asikainen, A., Bi, J., Depledge, M., van den Elshout, S., et al., 2016. Environ. Health 15 (Suppl 1), 25. http://dx.doi.org/10.1186/s12940-016-0097-0.
Silberzahn, R., Uhlmann, E.L., 2015. Crowdsourced research: many hands make tight work. Nature 526, 189–191. http://doi.org/10.1038/526189a.
Stanke, C., Kerac, M., Prudhomme, C., Medlock, J., Murry, V., 2013. Health effects of drought: a systemic review of the evidence. PLOS Curr. Disasters 1. Edition. http://dx.doi.org/10.1371/currents.dis.7a2cee9e980f91ad7697b570bcc4b004.
Swan, M., 2013. The quantified self: fundamental disruption in big data science and biological discovery. Big Data 1 (2), 85–99. http://dx.doi.org/10.1089/big.2012.0002.
UN Department of Economic and Social Affairs, 2013. World Population Ageing 2013 ST/ESA/SER.A/348 .
UN Department of Economic and Social Affairs, 2014. World urbanization prospects. In: The 2014 Revision, Highlights (ST/ESA/SER.A/3), .
UN Department of Economic and Social Affairs, 2015. World population prospects Volume II: Demographic Profiles (ST/ESA/SER.A/380) .
UNEP, 2013. Global chemicals outlook: towards sound management of chemicals. Available. http://www.unep.org/chemicalsandwaste/Portals/9/Mainstreaming/GCO/Rapport_GCO_calibri_greendot_20131211_web.pdf (accessed Aug 15 2016).
Vineis, P., Chadeau-Hyam, M., Gmuender, H., et al., 2016. The exposome in practice: Designof the EXPOsOMICS project. Int. J. Hyg. Environ. Health. http://dx.doi.org/10.1016/j.ijheh.2016.08.001.
Vrijheid, M., Slama, R., Robinson, O., Chatzi, L., Coen, M., van den Hazel, P., et al., 2014. The human early-life exposome (HELIX): project rationale and design. Environ. Health Perspect. 122, 535–544.
Wang, C., Chen, M.H., Schifano, E., Wu, J., Yan, J., 2015a. Statistical methods and computing for big data. http://arxiv.org/abs/1502.07989.
Wang, T.W., Vermeulen, R.C.H., Hu, W., Liu, G., Xiao, X., Alekseyev, Y., et al., 2015b. Gene-expression profiling of buccal epithelium among non-smoking women exposed to household air pollution from smoky coal. Carcinogenesis 36, 1494–1501.
Watts, N., Adger, W.N., Agnolucci, P., Blackstock, J., Byass, P., Cai, W., et al., 2015. Health and climate change: policy responses to protect public health. Lancet 386, 1861–1914.
Weisskopf, M.G., Sparrow, D., Hu, H., Power, M.C., 2015. Biased exposure – health effect estimates from selection in cohort studies: are environmental studies at particular risk? Environ. Health Perspect. 123 (11), 1113–1122.
Whitmee, S., Haines, A., Beyrer, C., Boltz, F., Capon, A.G., Ferreira, B., et al., 2015. The Rockefeller Foundation – Lancet Commission on planetary health safeguarding human health in the Anthropocene epoch: report of the Rockefeller Foundation – Lancet Commission on. Lancet 386, 1973–2028.
WHO, 2009. Protecting Health From Climate Change. ISBN 978 92 4 15981. Available. http://apps.who.int/iris/bitstream/10665/44133/1/9789241598187_eng.pdf

[accessed 15 July 2016]).
WHO, 2014. World Health Statistics. Available:. http://www.who.int/gho/publications/world_health_statistics/EN_WHS2014_Part3.pdf?ua=1 (accessed 15 July 2016).
WHO, & NIH, 2011. Global health and aging. Available:. http://www.who.int/ageing/publications/global_health.pdf (accessed 15 July 2016).
Wild, C.P., 2012. The exposome: from concept to utility. Int. J. Epidemiol. 41 (1), 24–32. http://dx.doi.org/10.1093/ije/dyr236.
Willett, W.C., 2002. Balancing life-style and genomics research for disease prevention. Science 296 (5568), 695–698.