Changing Contexts in Epidemiologic Research
– Thoughts of Young Epidemiologists on Major Challenges for the Next Decades

Ana Azevedo, Leda Chatzi, Tobias Pischon and Lorenzo Richiardi*

1 University of Porto Medical School & Institute of Public Health of the University of Porto, Portugal
2 Department of Social Medicine, Faculty of Medicine, University of Crete, Greece
3 Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Germany
4 Cancer Epidemiology Unit, CeRMS and CPO-Piemonte, University of Turin, Italy

1. Introduction

Since the second half of the last century, epidemiology has made key contributions to the identification of the causes of common chronic diseases, from cancers to cardiovascular and respiratory diseases, paving the way to effective prevention measures. In the process sophisticated epidemiologic methodologies have been developed for both observational and interventional studies in humans that have become the basis of “evidence-based” medicine and public health.
Two major developments have changed the scientific and societal contexts in which epidemiology operates: the extraordinary advances in molecular genetics, cell and developmental biology and a shift towards social philosophies in which individual values largely dominate over collective values. These developments affect on one side the aims, methods and contents of epidemiology as a research approach to health and disease and on the other side the public health perspective that confers a practical value to epidemiology.

Within this frame, a growing number of young epidemiologists, whose professional life will project over the next thirty or forty years, are involved in research, as indicated by the increasing number of epidemiology publications in the peer-reviewed literature. Less visible are, however, their activities concerning the medium and long term orientation and evolution of epidemiology, its role within public health and medicine and, ultimately, its ability to make a real difference to population health.

European Young Epidemiologists (EYE) is a network of young epidemiologists within the International Epidemiological Association – European Epidemiology Federation (IEA-EEF), founded in 2004, after the European Congress of Epidemiology. EYE aims to establish contact among young epidemiologists in Europe in order to facilitate future collaboration in scientific research, to engage in the development of epidemiological research methods, to foster the appropriate use of epidemiological research in the domains of public health and clinical medicine, and mainly to discuss and intervene in the future of epidemiologic research.

In 2011, a worldwide group of Early Career Researchers (ECR group) emerged within the International Epidemiological Association with representatives from all continents, and the chairman of EYE is the European representative at the worldwide group. This initiative reflects the felt need to network, promoting the discussion on what the future of epidemiology is and what it should be, by putting emerging epidemiologists’ voices on the map about how to make health research work towards scientific and societal development, ultimately contributing to improve populations’ health.

In this context, it seemed timely to us to share this chapter, which summarizes major topics that emerged from a 3-day workshop held in Turin, Italy, in 2008, to explore and debate the long term orientation and evolution of epidemiology. The workshop was organized by the European Educational Programme in Epidemiology in collaboration with the International Epidemiological Association and European Young Epidemiologists group and counted on the participation of 14 early career epidemiologists from different European countries and 7 experienced epidemiologists as discussants. This text does not intend to resuscitate a debate on the identity and role of epidemiology as a discipline, much less to offer a solution to these fundamental questions, but instead simply asks some questions about how epidemiology will respond to what we see as two major developments, as identified above - the inexorable rise of molecular biology and the shift from collectivist to individualist philosophy, which resonates with epidemiology's role as the basic science serving public health.

2. Development of epidemiology and its role in research

The importance of epidemiology as a scientific discipline has been steadily increasing over the past decades. In etiologic research this is partly driven by the intention to disentangle
the independent effects of risk factors for the predominant chronic diseases in Western civilizations, including cardiovascular diseases and cancer. These complex diseases are caused by many different factors, including genetic and non-genetic (diet, lifestyle, occupation, environment), where each factor is usually related to only small changes in risk. Consequently, the identification of risk factors usually requires the study of large samples with sophisticated analytic techniques, making this the prototype of epidemiologic research.

Epidemiology has evolved into several different subdisciplines which are focused on specific areas of research, such as cardiovascular epidemiology, cancer epidemiology, genetic epidemiology, clinical epidemiology or nutritional epidemiology. What is noteworthy, though, is that the advancement of epidemiology often seems to be driven by developments in these other specific areas of research (and, equally important, by researchers in these other disciplines), rather than by epidemiology itself (or epidemiologists themselves). For example, the search for genetic variants that may be associated with increased health risk has led to the creation of large databases and to the conduct of genome wide association studies (GWAS). Partly because of the large number of variants examined, GWAS have a high risk of providing false-positive results. This has led to discussions and suggestions about how to conduct and interpret results of such studies in the fields of genetics and genetic epidemiology (Hattersley & McCarthy, 2005), although the question of how to appropriately deal with false-positive results is essentially a genuine general question in epidemiology. Another example comes from the field of clinical epidemiology, where a majority of pertinent studies are performed by clinicians, although the questions addressed in these studies are mostly those that traditionally fall in the field of epidemiology.

One concern that applies to epidemiology as well as to other areas of scientific research is that advancements may in some instances simply be driven by the pure availability of new technologies or by advances in existing technology rather than by research questions pertinent to the area of research. An example is the development of various “-omics” technologies, which provides promising tools to allow large-scale biomarker studies, including discovery-oriented as well as hypothesis-testing investigations (Vineis & Perera, 2007). However, there may also be an inherent risk, i.e. that the research agenda is dictated by the ability to have novel or advanced technologies instead of having sound scientific questions or hypotheses. Thus, epidemiology may be “vulnerable” to the focus on novel technologies by a tendency to “enrich” studies with these technologies in an attempt to obtain higher impact without critical reflections about their usefulness. Importantly, a fact that may be neglected is that the use of novel technologies is often vulnerable to similar shortcomings as are the more traditional approaches, and may even add complexity to the interpretation of results. For example, with regard to the use of biomarker technologies it was succinctly cautioned that “biochemical measures are almost always subject to the same problems of misclassification and bias [as answers provided by humans and their interviewers]” (Hunter, 1998). Also, it must be stressed that several of the major breakthroughs in epidemiology, including case-control and cohort studies that led to the discovery of the role of smoking as the major risk factor for lung cancer, serum cholesterol and smoking as risk factors for coronary heart disease and folate deficiency as a determinant of neural tube defects, in fact came from a rather “old-fashioned”, black box approach (Susser & Susser, 1996). Although it may seem self-evident, it is important to reiterate that the research agenda in epidemiology should be driven by questions that primarily address topics that fall within the
definition of epidemiology, rather than by technical details of related disciplines. Taking molecular epidemiology as an example, McMichael stated in an editorial more than 10 years ago, “...we do not need a new ‘molecular’ subdiscipline, with an inevitable inbuilt tendency to reductionism. Rather, we should critically incorporate the emerging array of molecular biologic measurements into mainstream epidemiologic research and thus broaden its scope. Good science will come from a synthesis that transcends disciplines and techniques” (McMichael, 1994). A successful example of the exploitation of genetic technology by epidemiologists to study exposures that are difficult to measure is Mendelian randomization, a method of direct value to understanding the environmental causes of common diseases using genetic variants as proxies (Davey Smith & Ebrahim, 2003).

As indicated further in the next paragraphs, the future will probably bring the need to create and analyze even larger databases, implementing new technologies, and unraveling the complex interplay between environmental and genetic factors in disease etiology. What role will epidemiologists play in such endeavors? What is the relationship of epidemiology to other disciplines (genetics, clinical medicine, etc.) in the context of the creation and analysis of such databases? Will epidemiology be a method and will epidemiologists be the database managers? Or will epidemiology be a field that takes a leading role in shaping the research agenda? These are important questions which epidemiologists have to face for the future.

By providing their specific expertise epidemiologists need to make sure that they form an essential part of that process. This means that they need to be integrated in all parts of research, including the formulation of hypotheses, development of study designs, establishment and conduction of studies, analysis and interpretation of data, and translation into public health settings. For this reason it is of course essential for epidemiologists to have detailed knowledge of their areas of interest. However, while the creation of subdisciplines is an enrichment of the field of epidemiology, it is important to keep in mind the global aim of epidemiology, that is, to study “the occurrence and distribution of health-related states or events in specified populations, including the study of the determinants influencing such states, and the application of this knowledge to control the health problems” (Porta, 2008); nothing more, nothing less.

3. The research agenda of an epidemiologist

The question that we would like to address here is: "What are the determinants of the choice of our area of research in epidemiology?"; in other words: “Based on which criteria do epidemiologists decide on which research to follow?”. Even when a researcher has a complete independent status, the choice is the result of several forces and not restricted to the appreciation of which are the best scientific questions. In addition to scientific curiosity and public health relevance, many other factors have an implicit or explicit role.

Previous research experiences have great impact on our own research agenda. Changing one’s own specific area of research can be challenging, not only because of the need of new skills and knowledge but also because of the lack of national and international recognition and networking in the new research area, with consequent difficulties in being involved in collaborative research and having access to funding. Thus a change in area of research cannot be achieved in short time, while it needs long-term programming and a supportive research environment and infrastructure.
The research environment, the interaction with colleagues and their expertises and the facilities available at the research institute are obvious strong determinants of the research agenda.

Facilities also include the availability of large databases, an issue that will be discussed in the next section. Here we emphasize that the availability of administrative large databases has increased dramatically the opportunities for epidemiological research. Sometimes, however, the availability of data may also shape the research questions. As opposed to already available databases, collection of new data can be specifically targeted at emerging research hypotheses, but it may be hampered by cost and organisational constraints. Often, we favour a research question that can be answered using already available data as opposed to a research question that needs time- and resource-intensive studies. This approach may allow the risk of testing hypotheses with lower a priori likelihood of providing a consistent answer that is reproduced in other studies, standing the test of time; these hypotheses would otherwise not be approached, at least not immediately, thus possibly creating opportunities for discovery (Vandenbroucke, 2008). It is difficult to know what combination of the two approaches, use of available data and new collections, maximizes the possibilities of progress in scientific knowledge.

Hot topics are more likely to be published in more important journals, which, in turn, enhance the opportunities to reach the scientific community as well as lay people through the media. More important journals have also higher impact factor, that, although being criticized (Hernán, 2008), is still affecting researchers’ careers and access to funding. For their scientific and public health relevance as well for the reasons just described, hot topics are more likely to stimulate new research and to be considered as a research priority. This can translate in fast scientific progress and public health impact but, on the other hand, this process can divert resources and efforts from new developing fields.

A new field can emerge only if funding agencies are giving it adequate support. Indeed, funding agencies have a central role in shaping the research agenda, and, therefore, the transparency of their selection process is a fundamental issue. However, even a transparent selection that is strictly based on quality, public health implications and scientific relevance of the submitted projects does not limit the influence of the agencies. Often, funding agencies open specific calls for research aiming at a priori decided objectives. We feel that the extent of the role of public and private funding agencies in shaping research agenda should be measured and monitored over time, including an assessment of the process that leads to the definition of the specific calls and the actual societal impact of funded projects.

We started this section by considering an epidemiologist who has complete independence. The issue of independence has been studied and discussed in the epidemiological literature at length, mainly with reference to influences from the industry and, to a lesser extent, from governments (Pearce, 2008). Even assuming independence, however, we are aware of the fact that we all have a priori beliefs, we receive a salary from an institution or a funding agency, and we live in a community. It is widely accepted that epidemiologists should aim at giving priority to the research questions with the highest scientific interest and/or public health impact. This is however not a trivial task and we should recognise that the decision on what to study is affected by a large number of factors, many of which are not under our direct control.
4. Emerging opportunities and challenges in epidemiology: Large databases and use of secondary data

The twenty-first century undoubtedly provides new horizons regarding the availability and use of data sources. In the last decade a growing number of public databases for depositing data have emerged. Much of the impetus for this growing trend was given by the paradigm shift we witnessed in genetic research, which has moved from a candidate gene approach focused on few genes to GWAS, which require multifaceted linked databases of larger populations (Ioannidis et al., 2006; Kaiser, 2002; Wylie & Mineau, 2003). Although less common, similar trends in data storing and sharing occurred in other areas of epidemiology as well. One example is the Pharmacogenomics Knowledge of Base of PharmGKB (www.pharmgkb.org) that was established to store, manage and make available molecular data in addition to phenotype data obtained from pharmacogenetic studies. In the field of classical epidemiology, multi-centric collaborative studies and pooled analyses are becoming more and more common. Moreover, systematic reviews and meta-analyses try to integrate and synthesize existing research studies in an attempt to derive new information by quantitative statistical analysis. By examining the totality of data available about an issue, systematic review can identify inconsistencies in existing data and point to areas of research needed, reduce the potential for erroneous findings occurring by chance, and more accurately define the benefit and possible adverse effects of management strategies.

We feel that future epidemiological research will benefit greatly from the exchange of ideas between researchers and across disciplines/subdisciplines. This not only refers to concrete research results but also to approaches to the study of new areas. Existing studies could establish efficient routes of communication and co-ordination that allow a quick and detailed identification and promotion of common research areas. New studies could add protocols designed for specific purposes, preferably specialized rather than general, and study selected populations of special interest. A collaborative basis may in certain areas of research increase statistical power, ensure efficient design with large study populations, allow geographical comparisons and the replication of results, and give the possibility to study sub-groups or rare exposures (a crucial aspect of epidemiology) (Kogevinas et al., 2004).

Questions about ownership, custody and rights of access to data are major issues and determine restrictions to data sharing and collaborative research. These questions focus mainly on protection of privacy (the ability to control information about oneself) and confidentiality (the obligation of a second party to not reveal private information about an individual to a third party without the permission of the person concerned) (Willison, 1998). Confidentiality and privacy issues are emerging limiting factors (for both new data collection and use of available databases) that can have important effects on shaping research agenda and public health surveillance (Cuttini et al., 2009). At present, in many countries, legislations on confidentiality are defined with little consideration on their impact on medical and public health research, thus favouring personal privacy above societal benefits. The four principles of protection of a research participant are autonomy (self-determination), beneficence (maximal benefit), nonmalfeasance (minimal harm), and justice (distribution of benefits and harms across groups in society) (National Commission for Protection of Human Subjects of Biomedical and Behavioral Research, 1979). Although these principles focused on experimental studies in the past, it is essential that we follow
established ethical guidelines also in observational studies that are perceived to have minimal harm. In the past, these issues have primarily been raised with regard to clinical trials where the intervention itself may do harm to the research subject. In observational studies, however, the concern about harm is not so much about the fact that the study procedures may do harm to the research subject (which is usually minimal because of the observational nature) but more about the fact that the results of that research may (indirectly) harm the participants or a group thereof. For example, results of a genetic study may reveal that individuals with a certain genetic variation may have a higher risk of disease. Should researchers report these results to their study subjects? If yes, then such reporting could harm the self-determination of these subjects because they may not have asked for that specific test. If not, then the researcher may withhold important information from that person. If results with potential clinical significance are delivered to individual participants, the communication should be made in close collaboration with clinicians who should be part of the research group from the beginning of the project. In addition, if researchers decide to disclose the results, participants should have the opportunity at the time of enrolment to give their consent to receive information about incidental findings or not, and should receive explanations on how incidental information will be handled. Often these two requirements are not met or are unfeasible in a specific research project. Partly based on these concerns, some countries already adopted new laws or regulations, such as the Genetic Information Nondiscrimination Act in the United States in the year 2008 (Hudson et al., 2008).

As large electronic databases have been developed, several management models have been designed [e.g., the RGE (Resource for Genetic and Epidemiological Research) model, the Sweeney’s model, the deCODE Genetics model and others] focusing on confidentiality versus research use, as well as public versus private access (Wylie & Mineau, 2003). Individual rights of subjects must be respected at all times, but should not be misused by data collecting institutions as an argument to restrict access of other researchers. A balance between individual rights to privacy and the societal benefit of research must be established (Bergmann et al., 2008).

Another important issue when examining large databases is the frequent lack of explicit reports on the methods followed for the collection of the data from different sources, the completeness of this information, and a discussion of limitations of the data source. This may also be driven by the strict space limits of most journals as investigators may have had appropriately described everything in the methods section but word count limitations led to the deletion of this information. There are examples of large collaborative studies where all the methods and quality have been specified and assured (Tunstall Pedoe, 2003). There is a crucial need for researchers and journal editors to become aware that guidelines have been developed on how to conduct and how to report results of epidemiologic studies (International Epidemiological Association, 2007; von Elm et al., 2007).

The next step will be to enhance the availability of methods for easily depositing data and to provide tools for ensuring the sustainability of the databases. Large databases may benefit from widely available electronic search tools listing available studies on a specific topic and they should encompass both published and deposited data. A research environment that promotes and rewards by publishing only results that reach statistical significance is likely to foster data dredging and will create a distorted literature with very low credibility.
(Easterbrook et al., 1991; Ioannidis, 2005). The scientific community will also have to discuss issues of authorship, data property, and funding of secondary analyses.

The study of demographic, genetic, medical and environmental data from different populations may create an exciting and promising approach to identify the causes of common diseases and create effective preventive measures. “If you have large, accurate data sets on the health and death of human beings, what else do you need to improve the health of the public other than sound scientific method, cautious inference and a dialogue between science and policy?” (Coleman, 2007). Our knowledge of health and disease will certainly be greatly enhanced when the use of this immense amount of information is made available through the application of solid epidemiological principles. We are aware that there are problems to solve and agreements to reach within the field of large databases and use of secondary data. In addition, large databases and secondary analyses may not be useful to answer all new research questions, but they may be a (powerful) tool for epidemiological research.

5. Epidemiology and society: How each influences the other

Epidemiology tells us what we want to know about the human condition and, often, how it might be improved, in a way which no other science can offer (Coleman, 2007). This is a great challenge and a major reason why we find it so attractive and intellectually rewarding.

Throughout history, society has conditioned and channelled science. Societal reaction also influences the translation of epidemiology into public health. Many of the 20th century beliefs regarding the relation between epidemiology and society turned out to be only half-truths: 1) epidemiology would lead to prevention, 2) prevention was better than cure, 3) social justice would be achieved through prevention and 4) epidemiology would pervade clinical medicine and change its practice. We now recognize that success in epidemiology has not necessarily implied public health achievements (e.g. evidence on tobacco vs. economic interests) and health inequalities tend to increase instead of decrease.

The present loss of credibility before the society (and other fields of science) regarding risk factor epidemiology is partly a consequence of a reductionist view, i.e., a focus on associations between a single exposure and a single outcome, which frequently originates inconsistent messages (the same exposure may be publicized either as risk or protective factor on different adverse outcomes). Also, conflicting results regarding the same association might raise the question of how much evidence is needed to intervene or to advocate intervention (Taubes, 1995). Publication of small amounts of information without considering implications contributes to incomplete knowledge and in our view reflects some degree of irresponsibility. Publication drive may result in objective dishonesty that must be fought against. Introspection should be carried out before publication: are we honestly convinced by our findings?

Etiological epidemiology has mostly been looking at individual susceptibility and the distribution of disease in the population has been undervalued. The growing emphasis on genetic/molecular research contributes to direct epidemiology towards individual-based prevention as opposed to population level approaches. Concern with individual susceptibility has neglected the distribution of disease in the population, leading to the “type III error” – a good study to answer the wrong question. While an increased interplay
between biotechnology, infrastructures and methods may be the future of epidemiologic research, translational research must be promoted, starting from the population and responding to its needs, with special attention being required towards understudied groups (e.g. migrants).

Political stability is an important basis for public health. Inequalities in health and research between countries, even within Europe, emphasize the need for a) one epidemiology for all societies in the 21st century, b) more quality research from less rich countries, c) stronger political will to translate evidence into action.

The reinforcement of epidemiologists’ professional image with society in general is needed. The importance given to individual values such as the right to privacy has risen barriers to research that in our view do not benefit society as a whole while in fact the risk of disrespect for individual rights is smaller than its theoretical maximum. There is the need to distinguish between the risk to personal autonomy from the use of identifiable data without consent to select a given individual for prurient interest or unauthorized disclosure (moving from population data to the individual) and the far smaller risk posed by aggregating individual data for research in order to draw general conclusions about society (from individual data to the population) (Coleman, 2007). Striking the right balance between the confidentiality of identifiable health data and the need for medical research to improve public health is now an issue in many countries (Coleman et al., 2003). Though it is not necessarily straightforward where the line should be drawn, the societal pendulum needs to swing back towards the collective responsibility for medical research and public health surveillance. Current regulatory climate risks to refrain the scientific community from using available data to control health problems and improve population health.

We feel the need for a strengthening of the link between epidemiologic research and society, in order to translate findings into the effective improvement of population health. Part of this process should be the reinforcement of epidemiologists’ professional image in the society in general to win its trust.

6. Conclusion

Research has been strongly influenced by a random and passive intersection between biotechnology, infrastructures and available methods. Young epidemiologists must reinforce their knowledge on the substantive issues they are researching and promote an active interaction between biology and society. Translational research is needed to use relevant laboratory research resources in population-based studies and to make the results of epidemiological studies useful to an individualized and predictive medical practice.

Professionals need to be prepared to collate data. Questions about ownership, custody and rights of access to data are major and determine restrictions to research. Individual rights of subjects must be respected at all times, but should not be misused by institutions that collected data as an argument to restrict access of other researchers. More than new information, we need to use the information we already have. A balance between individual rights to privacy and the societal benefit of research must be encountered.

In order to gain the possibility of playing a more active role in their research agenda, epidemiologists must improve their communication skills, both regarding risk
communication to the population and scientific dialogue with other researchers and clinicians. Also, they need to conquer a position in funding agencies and as consultants for policy makers, and be available for these tasks over time.

The need to reinforce the professional image of epidemiologists could be met by as good a formal education as possible along with good epidemiologic practices. Epidemiological expertise will continue to be required for the attempt to set rational priorities for the control of disease and health promotion. This challenge is as breathtaking as we need to keep us on track to contribute to design the future of epidemiology.

7. Acknowledgment

This work was supported by Compagnia di S. Paolo, to whom the authors gratefully acknowledge all the material conditions for the workshop “Epidemiology in the new century: a perspective from the young European epidemiologists”, held in Turin, Italy, in May 2008.

The liveliness of the discussions at the workshop and its output would not have been possible without the generous contribution of the senior epidemiologists who attended – Shah Ebrahim, Hans-Werner Hense, Franco Merletti, Jorn Olsen, Susanna Sans, Rodolfo Saracci, Paolo Vineis. The authors want to particularly thank Rodolfo Saracci for his intellectual input, as well as his initiative and enthusiasm in the organisation of this event, which were a sine qua non condition for the workshop and all the outputs thereafter.

8. References


This special issue resulted from the invitation made to selected authors to contribute with an overview of a specific subject of their choice, and is based on a collection of papers chosen to exemplify some of the interests, uses and views of the epidemiology across different areas of research and practice. Rather than the comprehensiveness and coherence of a conventional textbook, readers will find a set of independent chapters, each of them of a great interest in their own specialized areas within epidemiology. Taken together, they illustrate the contrast between the attempt to extend the limits of applicability of epidemiological research, and the “regular” scientific activity in this field or an applied epidemiology. Epidemiologists with different levels of expertise and interests will be able to find informative and inspiring readings among the chapters of this book.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: