

# Realizing the promise of population biobanks: a new model for translation

Madeleine J. Murtagh · Ipek Demir ·  
Jennifer R. Harris · Paul R. Burton

Received: 18 May 2011 / Accepted: 5 June 2011 / Published online: 25 June 2011  
© The Author(s) 2011. This article is published with open access at Springerlink.com

**Abstract** The *promise* of science lies in expectations of its benefits to societies and is matched by expectations of the realisation of the significant public investment in that science. In this paper, we undertake a methodological analysis of the science of biobanking and a sociological analysis of translational research in relation to biobanking. Part of global and local endeavours to translate raw biomedical evidence into practice, biobanks aim to provide a platform for generating new scientific knowledge to inform development of new policies, systems and interventions to enhance the public's health. Effectively translating scientific knowledge into routine practice, however, involves more than good science. Although biobanks undoubtedly provide a fundamental resource for both clinical and public health practice, their potentiating ontology—that their outputs are perpetually a *promise* of scientific knowledge generation—

renders translation rather less straightforward than drug discovery and treatment implementation. Biobanking science, therefore, provides a perfect counterpoint against which to test the bounds of translational research. We argue that translational research is a contextual and cumulative process: one that is necessarily dynamic and interactive and involves multiple actors. We propose a new multidimensional model of translational research which enables us to imagine a new paradigm: one that takes us from bench to bedside to backyard and beyond, that is, attentive to the social and political context of translational science, and is cognisant of *all* the players in that process be they researchers, health professionals, policy makers, industry representatives, members of the public or research participants, amongst others.

---

M. J. Murtagh (✉) · P. R. Burton  
Department of Health Sciences, University of Leicester,  
Adrian Building, University Road, Leicester LE1 7RH, UK  
e-mail: mm399@le.ac.uk

I. Demir  
Department of Sociology, University of Leicester,  
University Road, Leicester LE1 7RH, UK

J. R. Harris  
Division of Epidemiology, The Norwegian Institute of Public  
Health, Oslo, Norway

P. R. Burton  
Department of Genetics, University of Leicester,  
University Road, Leicester LE1 7RH, UK

P. R. Burton  
Department of Epidemiology, Biostatistics and Occupational  
Health, Faculty of Medicine, McGill University, Purvis Hall,  
1020 Pine Ave. West, Montreal, QC H3A 1A2, Canada

## Introduction

Expectations of scientific research run high. Internationally, funding bodies and government agencies have placed increasing emphasis on the outcomes and impact of scientific endeavour (Campbell et al. 2000; Cooksey et al. 2006; Craig et al. 2008; Zerhouni 2003). The *promise* of science lies in expectations of the benefits bioscience can bring to societies and is matched by expectations of the realisation of the significant public investment in that science. In health research, the transformation of bioscience into societal benefit is described in terms of translation. First reported as such in the medical literature in 1993 the concept of translational research has a much longer history (Antoine 1991). The increasing interest, especially recently, is evident in even a cursory search of PubMed where citations including the term 'translational research' have increased nearly 200-fold since 1993, with more than half of these

citations occurring in the past 2 years. Likewise, the use of the term ‘bench to bedside’ to describe the transformation of bioscience into therapeutic practice, seen as early as 1985 (Merz 1985), has exploded across the biosciences. In addition, therein lies a conundrum for translational research in biobanking.

Population biobanks represent an ontologically distinct approach to knowledge generation when compared with the dominant scientific forms envisaged by translational research. Population biobanks merge the promise of genomics with foundational public health science, including classical epidemiology (Khoury et al. 2009a, b), principally in the form of population-based cohort studies (Burton et al. 2010; P3G Observatory 2009). Biobanking science is implicitly oriented to the *population focus* of its historical epidemiological forebears as well as to the *individualist focus* of contemporary genomic medicine. Biobanking steers away from the hypothesis testing that is fundamental to much basic and applied bioscience; i.e. genome-wide association studies (GWAS) and clinical trials, respectively. Instead, it embraces methods of knowledge generation which prepare us to ask questions we may not yet be able to formulate. Extensive work has already been undertaken (Knoppers et al. 2008) to refine the design (Burton et al. 2009; ISBER 2005; Moore et al. 2011; Wallace et al. 2009), management (Borugian et al. 2010; Litton et al. 2003; Peakman and Elliott 2008; Yuille et al. 2008) and harmonisation (Fortier et al. 2010; Stover et al. 2010; Wichmann et al. 2011) of biobanking platforms and to promote and facilitate liberal data access (P3G Consortium et al. 2009; Wolfson et al. 2010).

However, biobanks and data are not ends in themselves. Part of global and local endeavours to translate raw biomedical evidence into practice, their purpose is to provide a platform—composed of high quality data and samples—that can later provide a firm foundation for generating new scientific knowledge to inform development of new policies, systems and interventions to enhance the public’s health (Davey Smith et al. 2005; Khoury et al. 2009a). Effectively translating scientific knowledge into routine practice, however, involves more than good science. It is a complex social process that necessarily involves scientists, health professionals, policy makers, funders, industry and members of the public (including research participants); i.e. it is about the people not just the science. Moreover, the mechanisms of successful translation are not well enough understood. Drawing on a range of disciplinary perspectives, including biomedical and public health science, social science and philosophy, we overview the scientific logics underpinning the creation of biobanks and describe what translation might mean for population biobanking research. We argue that if the future utility of population biobanking is to be optimised we must explore what we

should be doing *now* to ensure extension of the transdisciplinary collaborations to include, for example, health policy makers so that biobanking resources are configured with translational aims in mind. This will help ensure that biobanks can best contribute to the evidence-base for health policy decision-making, health-care commissioning and the implementation of new initiatives in practice.

### Population biobanking science

The generation of scientific knowledge involves the asking and answering of questions within constraints imposed by contemporaneous needs, knowledge and technology. Until recently, most definitive answers in health science pertained to factors with relatively large effects—e.g. Lind’s demonstration that citrus fruit prevents scurvy (Kirch 2008); the causal role of smoking in many diseases (World Health Organisation 2008); the impact of monogenic genetic variants causing major diseases such as Huntington’s chorea and cystic fibrosis (Carter 1977). Recent years, however, have seen a fundamental change in the nature of biomedical investigation. Modern bioscience reflects the disease priorities of *contemporary* societies—particularly those of the ‘High Income Countries’ that primarily fund that research. There is therefore a manifest focus on the aetiological architecture of the common chronic diseases that affect these societies (Merikangas and Risch 2003; Pritchard and Cox 2002). The aim is to understand, and ultimately intervene in, the complex causal pathways via which genes and environmental/lifestyle/social determinants act alone and in interaction to cause diseases and to influence how those diseases progress once they have developed (Khoury et al. 2009b; Merikangas and Risch 2003).

This new challenge demands a comprehensive exploration of gene networks, of much weaker main effects (Manolio et al. 2008; Merikangas and Risch 2003; Wellcome Trust Case Control Consortium 2007) and of interactions between genetic, environmental, behavioural and social factors (Manolio et al. 2006). Typically, the aim is to identify and quantify the impact of a number of small effects (genetic, socio-environmental or a combination) on a given disease or health-related trait, but to do this the effects of the determinants of interest must first be disentangled from an obscuring cacophony of aetiological noise resulting from many other causal determinants. The modern search for the risk factors of complex diseases has often been compared to finding a needle in a hay-stack (Moore et al. 2010), but this simile should probably be extended to note that the needle itself is made of straw.

The change in focus from investigating risk factors with relatively large effects to explore networks of risk factors

with weak effects requires a radically different approach. This is because many of the effects are *very* weak: for example, although the relative risk of lung cancer in smokers versus non-smokers is of the order 9.0 (Doll and Hill 1952, 1964), most known associations between genetic variants and chronic diseases reflect weak effects with typical allelic odds ratios in the range of 1.1–1.4 (Burton et al. 2009; Spencer et al. 2009; Zondervan and Cardon 2007). Critically, effects as weak as this not only provide an obvious challenge from the perspective of statistical power (Burton et al. 2009) but, in addition, they can easily be created or concealed by confounding or reverse causality. In fact, these latter concerns led Gary Taubes to claim that observational epidemiology had already reached its inferential limits (Taubes 1995).

There is no doubt that useful scientific inferences can be generated in such settings, but the obvious route to new knowledge would involve well-designed experiments: carefully controlled experiments in the laboratory setting, or studies involving randomised allocation of intervention as in a conventional clinical trial. Such experiments can negate confounding. However, many research questions in human health are simply not amenable to an experimental approach. It is, therefore, fortunate that despite the very real concerns raised by Taubes (1995), useful inferences on weak effects *can* sometimes be based on observational studies.

Although genetic variants cannot be allocated randomly as part of an experiment—they *are* allocated randomly at conception (Davey Smith 2006; Davey Smith et al. 2005). Specifically, if one of your parents has two different alleles at a particular locus, you must receive one, and only one, of them (Burton et al. 2005). Furthermore, the particular allele that is actually transmitted is randomly determined in a manner that is entirely independent of the variants that are inherited at all other loci (from either parent), with the specific exception of those few loci that are in close linkage disequilibrium (LD) (Clayton and McKeigue 2001; Cordell and Clayton 2005) on the same chromosome. This process forms the basis of what is often known as Mendelian randomisation (MR) (Davey Smith 2006; Davey Smith and Ebrahim 2003; Sheehan et al. 2008) which has a number of profound implications: (i) genetic variants do not usually confound one another unless they *are* in close LD (here we ignore the inferential distortion that can arise from ancestral substructure in an imperfectly designed study; Devlin et al. 2001). (ii) Genetic variants are not generally associated with socio-environmental determinants (and cannot therefore confound them) unless a particular genetic variant has a biological mechanism that directly modifies exposure to a socio-environmental determinant of interest. In other words, which of two alleles is received at conception from one parent cannot possibly be related to whether an individ-

ual later chooses to smoke—unless those two alleles have a different and direct impact on their predilection to smoke. In general, therefore, inferences on genetic determinants are not confounded by environmental factors and vice versa. (Clayton and McKeigue 2001; Davey Smith 2006; Davey Smith and Ebrahim 2003; Davey Smith et al. 2005; Sheehan et al. 2008). (iii) Interactions between socio-environmental determinants and genes are typically more robust to confounding than are the direct effects of the socio-environmental determinants themselves. For example, in relation to the role of a gene (G) and a socio-environmental determinant (SE) in causing a disease (D), serious confounding of the direct effect of SE could easily arise if confounding factor C was associated both with D and with SE and such confounding is extremely common because health-related behaviours tend to cluster (Davey Smith et al. 2005; Ebrahim et al. 2004): e.g. smokers often take less exercise, drink more alcohol, and take fewer vitamins. This is why it was argued that observational studies have a limited role in relation to studying weak effects (Taubes 1995). But, if scientific interest focuses on the *interaction* between G and SE (G:SE)—i.e. does the impact of SE vary differentially with genotype at G?—inferences are markedly less sensitive to conventional confounding as described above. It is true that serious inferential problems *will* arise if the nature of the G:SE *interaction* varies systematically with the level of C, but this is markedly less likely than the near ubiquitous clustering of health-related behaviours (Davey Smith et al. 2005; Ebrahim et al. 2004). (iv) Items i, ii and iii imply that weak effects associated with genetic variants or with interactions between genes and socio-environmental determinants are of considerably greater inferential value than would be the case in the absence of MR (Taubes 1995). However, Mendelian randomisation also provides a firm inferential foundation when scientific interest focuses on the direct effect of an environmental exposure alone. In fact this is the very purpose to which MR has classically been applied (Clayton and McKeigue 2001; Davey Smith 2006; Davey Smith and Ebrahim 2003; Davey Smith et al. 2005; Sheehan et al. 2008; Timpson et al. 2005). Specifically, if one is interested in whether factor SE protects against disease D, if one can identify a genetic variant (G\*) that has the same impact on an individual as increased (or decreased) exposure to SE, then demonstration that G\* is associated with D provides evidence that SE is causally associated with D. This is known as an “instrumental variable” approach (Davey Smith et al. 2005; Didelez and Sheehan 2007; Sheehan et al. 2008). Crucially, although this approach negates both confounding and reverse causality, a number of key assumptions that demand proper understanding of the underlying biology and mathematics must be met (Didelez and Sheehan 2007; Palmer et al. 2008; Sheehan et al. 2008).

In addition, if enough is known about the underlying biological, environmental and social determinants and if adequate resources can be invested in measuring relevant factors extremely carefully, it will sometimes be possible to control the analysis closely enough to enable inferences to be drawn that may reasonably be viewed as robust to confounding. Furthermore, if such an analysis is undertaken as part of a prospective cohort study (with initial exposure assessments being made whilst participants are still healthy) reverse causality is far less likely to be a serious problem. As we learn more and more about the underlying biology and social science, our ability to address confounding in ways that use prior knowledge in concert with sophisticated analysis based on the data from very comprehensive databases will inevitably increase.

If they are used cautiously and appropriately, it is therefore clear that large observational epidemiological studies *can* provide a useful basis for investigation of the causal architecture of complex diseases. However, it is also clear that these studies must be designed, set up and managed strategically (BBMRI 2010; Knoppers et al. 2008; Knoppers and Kent 2006). This need was reflected in the decision to include European Biobanking and Biomolecular Resources amongst the 36 designated priorities in the ESFRI Roadmap that represented the European Union's strategy for development of the infrastructure supporting pan-European science (European Strategy Forum on Research Infrastructures 2006). In effect, biobanks are now seen as the infrastructural equivalent of linear accelerators and telescope arrays in the physical sciences: that is, multi-purpose facilities representing broad-based platforms to support the scientific community as a whole in its asking and answering of key questions across the realm of bioscience. However, if their value is to be optimised, these facilities must have a number of pivotal characteristics:

1. They must provide access to data and biosamples relating to vast numbers of individuals (Burton et al. 2009; Collins 2004; Khoury 2004; Spencer et al. 2009; Zondervan and Cardon 2007). Case-control studies including *thousands* of cases are required even when interest focuses on the most straightforward situation: the detection of simple associations between single nucleotide polymorphisms (SNPs) (Burton et al. 2005) and a disease of interest (Burton et al. 2009; Spencer et al. 2009; Zondervan and Cardon 2007). Furthermore, when, as is likely, scientific emphasis moves on to focus on the study of gene-socio-environment and gene-gene interactions, and the exploration of causal pathways more comprehensively, *tens of thousands* of cases will often be required (Burton et al. 2009). Tens of thousands of participants can also be required to study a quantitative phenotype (e.g. measured blood

pressure), because allelic effect sizes may be as small as 1/10th of a standard deviation, or even less (Newton-Cheh et al. 2009a; Repapi et al. 2009). Finally, the use of an instrumental variable approach to address confounding increases the sample size requirement still further (Clayton and McKeigue 2001).

2. Data and samples must be collected under formal standard operating procedures (Fortier et al. 2010; ISBER 2005; Moore et al. 2011), and the assessment of disease status, biomarkers, physiological processes and social and environmental factors must be both accurate and precise. When aetiological effects are already weak, measurement error can substantially reduce statistical power yet further (Burton et al. 2009; Wong et al. 2003), and if errors are made in assessing confounders, real causal relationships may be obscured or artefacts created.
3. To achieve sample sizes of the magnitude required—i.e. involving enough participants that each have high quality data and samples of the nature required—it will often be necessary to pool data across multiple studies (Burton et al. 2009, 2010; Wichmann et al. 2011). This is well-exemplified by the large collaborative consortia that have been responsible for much of the recent progress in human population genomics (Burton et al. 2007; Easton et al. 2007; Frayling et al. 2007; Hindorf et al. 2009; Newton-Cheh et al. 2009a, b; Saxena et al. 2007; Scott et al. 2007; Stacey et al. 2007; Zeggini et al. 2007).

Collectively, these various requirements have led to the development of a new field of endeavour in bioscience that may be referred to as 'biobanking science'. The many professionals and researchers that work in this field include: laboratory scientists, clinical scientists, epidemiologists and statisticians, informaticians, ethico-legal experts and social scientists. Although they each have their own specialist areas of exploratory interest and expertise within the broader field of biomedical science, they share a *collective* focus on optimising the design, set up, harmonisation, networking and evaluation of the range of studies that are sometimes called biobanks: "*Organized collection[s] of human biological material* [e.g. blood, urine or extracted DNA] *and associated information stored for one or more research purposes*" (Burton et al. 2010). The studies that fall under this umbrella primarily represent population-based cohort studies, large case-control studies and tissue repositories, but some cross-sectional studies and randomised controlled trials qualify as biobanks too (Burton et al. 2010; P3G Observatory 2009). Crucially, although many of these studies were originally designed to support classical "phenotype to genotype" science (e.g. nested case-control studies), they also provide a platform for exposure-based

investigation including “genotype to phenotype” or “genotype-based” (Burton et al. 2010) studies. This latter approach uses exposure-based information (particularly, genotype data) to identify informative population-subsamples that, with consent, can be brought in for detailed additional investigation. Here, the vast size of modern cohort-based biobanks—sometimes viewed as being scientifically inefficient—provides a great strength: even a genotype that occurs as rarely as one in one thousand people, will be represented in five hundred participants in a population-based biobank containing 500,000 participants such as UK Biobank. In addition, when biobanks are effectively catalogued and harmonised (Burton et al. 2010; Fortier et al. 2010; P3G Observatory 2009), the available sampling frame for such studies can potentially be extended to several millions.

The structures and actors within biobanking science are reflective of what is described within social studies of science as Mode 2 knowledge production; that is, a new way of thinking about the generation of knowledge in science (Nowotny et al. 2001, 2006). Mode 2 knowledge production is carried out in the context of application (research translation), brings a heterogeneity of skills and expertise to the problem solving, and is transdisciplinary in nature (Nowotny et al. 2001). Transdisciplinarity is not simply another of the proliferation of terms for multidisciplinary, but rather points to the transgressive character of knowledge generation (Nowotny et al. 2001) that there is a co-evolution or co-production (Callon et al. 1992; Callon 1999, p. 69; Jasanoff 2004, p. 70; Geels 2005) of knowledge that transgresses disciplinary boundaries. Thus, it reflects the *intersection* and co-evolution of disciplines. Transdisciplinarity requires actors to engage actively in the messy and thorny work of overcoming difference and awkwardness whilst also drawing strength and value from the very existence of those differences (Demir 2008, 2011). Demir (2011) demonstrates how scientists achieve translation of their work across disciplinary borders by extending their own concepts and working language to incorporate new concepts; much as in second language acquisition. Through informal and formal exchanges, scientists in one discipline acquire, albeit in a limited way, the language, ideas and practices of another group (Demir 2011). Integration of these borrowed concepts and practices reshape the frameworks of the adopting discipline or field; which might, for example, facilitate harmonisation and data sharing. Within biobanking, there is clear evidence not only of transdisciplinary structures, but of the transdisciplinary generation of knowledge. An ongoing ethnographic study (Hammersley and Atkinson 2007) of the development of DataSHIELD (Wolfson et al. 2010) by authors of this paper (M.M., P.B.) demonstrates how the intersection of ethics, informatics and statistics is shaping the structure of the ensuing technology that is DataSHIELD.

The rapid progress of biobanking science over the last 5 years reflects a successful professional networking of biobanking scientists across the world that has been led by organisations, institutions and projects such as: P<sup>3</sup>G (Knoppers et al. 2008), PHOEBE (Wolfson et al. 2010), BBMRI (Wichmann et al. 2011), GEN2PHEN (2010), BioSHARE-eu (2011), ISBER (2005), and NCI (Moore et al. 2011). Crucially, the work carried out by these complementary and overlapping groups has, to date, focussed primarily on designing and setting up biobanks and on exploring a variety of approaches to networking and harmonisation (Borugian et al. 2010; Burton et al. 2009; Fortier et al. 2010; ISBER 2005; Litton et al. 2003; Moore et al. 2011; Peakman and Elliott 2008; Stover et al. 2010; Wallace et al. 2009; Wichmann et al. 2011; Yuille et al. 2008). However, the scientific focus is now entering a new phase in which the priorities are to provide liberal, but secure, access to data and samples (Pisani and AbouZahr 2010; Walport and Brest 2011), to ensure that biobanks are used widely and effectively, and to streamline pooled analysis involving multiple biobanks (Fortier et al. 2010; Wolfson et al. 2010).

The success of the large transdisciplinary organisations that comprise international biobanking is in no small part due to the fact that the science underpinning the successful construction and use of biobanks is philosophically very different from that underpinning hypothesis-based science. Although biobanks are undoubtedly a child of contemporary bioscience and are aimed primarily at facilitating new scientific discoveries, they are not hypothesis-driven per se. Rather, they may be seen as being part of a pre-competitive endeavour on the part of many bioscientists to advance and facilitate future hypothesis-driven science. In consequence, although biobanks are costly and may take many years to mature to full value, they should not be viewed as *competitors* to hypothesis-based science. Rather they should be viewed as enablers—i.e. an essential pre-requisite if hypothesis-based research is to be as successful at disentangling the causal structures underpinning complex diseases as it has historically been at revealing the impact of much stronger determinants of health and disease. In essence, biobanks will ensure a flexible yet valid platform for the as yet unpredicted (and unpredictable) bioscience questions of tomorrow. This not only impacts on the scientific design of biobanks, but it has major ethico-legal and social implications. How, for example, can we effectively ask participants to consent to answer detailed questionnaires, provide invasive tissue samples and agree to long-term monitoring of their health when even the scientists do not know how all of the data and samples will be used in the future? Of course, one possible answer to this, seemingly rhetorical, question is to ‘ask them’.

There can be little argument, but that the development of biobanking science over the first decade of the new

millennium has been remarkable. Although it may be of great interest and reward to the scientists involved, no part of bioscience can ever be an end in itself. Society invests heavily in scientific research related to medicine because it is assumed that this will ultimately improve the health of the community and of individuals within that community (Cooksey et al. 2006; Martin et al. 2008). The *raison d'être* of biomedical research is therefore to enhance public health through primary prevention, improved diagnosis and treatment, and the enhanced long-term control of chronic diseases. Such ends may be achieved via a wide range of different interventions at either the individual or population levels and biobanks can undoubtedly make a pivotal contribution to these aims. But if biobanks are to achieve optimal utility in the long term, it is critical that extensive thought is *now* put into how we should build on the successes of biobanking science to date, to ensure that biobanks are able to provide effective translational return in the future. The transdisciplinary collaborations must be extended to include professionals in public health, health policy making and health economics. In addition, we should consider the embedding of biobanks into health care systems. This will facilitate the enrichment of biobank databases with population-based and hospital-generated health event data. It will also enhance the future potential for biobank-driven translational activity; particularly, translation into population level public health. It is these issues that provide the focus for the remainder of this paper.

### Translational research: transforming data into knowledge and knowledge into practice

Demonstrating what has been described as a shift from *regimes of truth* to *regimes of hope* in clinical and bioscience (Moreira 2009; Moreira and Palladino 2005)—that is, the speculative *potential* of new therapeutic interventions—discussion of translational research begins in earnest with the NIH roadmap (US) (Zerhouni 2003), Cooksey report (UK) (Cooksey et al. 2006) and MRC framework for complex interventions (UK) (Campbell et al. 2000; Craig et al. 2008). The NIH describes translational research as including two areas of translation. “One is the process of applying discoveries generated during research in the laboratory, and in preclinical studies, to the development of trials and studies in humans. The second area of translation concerns research aimed at enhancing the adoption of best practices in the community” (NIH 2011). That is, *bench to bedside* and *research into practice*. Called T1 and T2 translational research, respectively, we can think of these endeavours as the processes and practices of transforming *data into knowledge* and *knowledge into practice*. The Cooksey report identified two gaps in research, “the first gap arises in

the translation of basic and clinical research into ideas and products”, i.e. at the bench; “the second gap relates to introducing those ideas and products into clinical practice”, i.e. at the bedside (Cooksey et al. 2006). Whilst the NIH *areas of translation* describe a transition from the bench to the community, *the backyard*, it is gaps in T1 translational research, from bench to bedside, that are the predominant focus of the translational research literature and funding (Woolf 2008). Thus, the long standing viewpoint of the biomedical lens, which more often includes the individual than the population, is maintained (Murtagh and Hepworth 1997; Thomson et al. 2005). Nonetheless, just *how* these translational transitions might be bridged has been the subject of much discussion, though there is a paucity of empirical research on these processes. What is clear, however, even in the seemingly most straightforward of translational processes, for example from biomarker to drug discovery to testing to adoption in routine clinical practice, is that it is not sufficient to simply throw the outputs of science “over the wall” (Cox 2011). That is, without aim, direction, purpose and without communication on either side between scientists, clinicians and others involved in disseminating and implementing that science. There is recognition within the bioscience community that translational research presents challenges and that the translational research model as initially described does not adequately account for the range of activity within bioscience or for the embeddedness of bioscience in and with other disciplinary, professional or public groupings and settings. Early appeals for the involvement of a range of disciplines envisaged an expansion to include biomedical and informatics scientists with bioscience (Zerhouni 2003, 2005). More recent expectations reflect what social studies of science describe as *communities of promise* (Martin et al. 2008) in which the articulation of clinical and biosciences are fundamental to the dynamics of translation and innovation and include the involvement of social, behavioural and ethico-legal disciplinary contributions (Khoury 2010; Khoury et al. 2009b, 2011; Zimmern and Brice 2009) as well as public and participant engagement in translational research, especially in the knowledge to practice phase (Armstrong et al. 2006; Graham and Tetroe 2007; Khoury et al. 2007). The importance of such contributions has been demonstrated empirically (Löwy 1996).

Elaborations of the NIH/Cooksey model have largely included disambiguation of the T2 phase. Westfall et al. (2007) argue for an expanded conception of the NIH translational research “roadmap” to ameliorate the “myriad detours, speed traps, roadblocks, and potholes [that] limit the movement of treatments from bench to practice” i.e. the limited external validity of randomised controlled trials, the diversity of health care practice efficiency and effectiveness, limited successful collaboration between researchers,

**Fig. 1** An expanded translational research continuum (1 term used by Burke et al. 2010; 2 derived from Green and Guyer 2011, 3 terms used by Khoury et al. 2007)

T1 Phase					T2 Phase		
Data to Knowledge					Knowledge to Practice		
Understandings of health and disease					Implementation and Improvement		
State of Population Health	Molecular	Cellular	Somatic	Social & Population	To the Patient	Into Practice	Embedding in routine practice
T0 <sup>1</sup>	T1.1 <sup>2</sup>	T1.2 <sup>2</sup>	T1.3 <sup>2</sup>	T1.4	T2 <sup>3</sup>	T3 <sup>3</sup>	T4 <sup>3</sup>
Translational Tools and Methodologies – facilitating knowledge generation and implementation							
T0 <sub>tm</sub>		T1 <sub>tm</sub>			T2 <sub>tm</sub>		

clinicians and patients and failure to address the identified needs of the community. They expand translational research to incorporate T1–T3 components: where T1 represents *translation to humans* of basic science in the form of preclinical and animal studies; T2 incorporates *translation to patients* of research through guidelines development, meta-analyses and systematic reviews; and, in T3 the *translation to practice* via dissemination and implementation research (Westfall et al. 2007). To address “mounting expectations” of genomic medicine Khoury and colleagues describe a T1–T4 continuum to “accelerate appropriate integration of human genome discoveries into health care and disease prevention” (Khoury et al. 2007). In the context of the hoary debates regarding population impact (public health) and individual outcomes (personalised medicine) of genomic discovery, Burke et al. (2010) have most recently added a T0 pre-scientific phase in a T0–T4 model to represent the context of population health needs. The T0 phase represents *assessment of the population health-disease burden* (Burke et al. 2010) which prompts (whether formally or informally) T1 translational research that aims to move *scientific discovery* into *candidate applications*. Evaluation of candidate applications (treatments and other health interventions) in T2 translational research assesses these applications and shapes development of evidence-based guidelines and policy. T3 translational research aims to move *evidence-based recommendations and policy* into health care practice via implementation, dissemination and diffusion research. Phase four, T4 translational research aims to evaluate the health impact of the *application in practice*. And thus the cycle begins again albeit, we hope, with a modified population health-disease burden.

Though not addressing the T1/T2 model explicitly, the recent review by Green and Guyer (2011) references a disarticulation of the T1 phase. Describing translation in genomics as moving from base pairs to the bedside, Green and Guyer (2011) add understanding of “the structure of genomes”, the “the biology of genomes” and “the biology of disease” to the T1 model. A molecular, cellular and somatic disease model implied by Green comprise what we

call the T1.1, T1.2, T1.3 components of the T1 phase (see Fig. 1). Added to this, we include T1.4, the understanding of *the social context and determinants of disease*: this may be in terms of the social construction and experience of the disease itself (cf. Murtagh and Hepworth 2003), understandings of the social context of disease to inform new interventions in the T2 setting (cf. Farnworth et al. 2008), or population level studies of disease determinants. As we state above, understanding of the social, behavioural and ethico-legal dynamics are fundamental to successful implementation and improvement science. Also missing from current models of translation is an explicit acknowledgment of the tools and methodologies that underpin translation of basic and applied research: for example, innovative theories, tools and methods for analysing data to maintain privacy and confidentiality (Wolfson et al. 2010); evaluating the appropriateness of interventions (Murtagh et al. 2007); enhancing the processes of knowledge translation and exchange (Graham and Tetroe 2007; Kitson AS 2009); improving practice (Dixon-Woods et al. 2001); implementing policy guidelines and new practice in complex organisations (May 2006; May et al. 2009) and other approaches of implementation and improvement science. We call these T<sub>tm</sub> or, T0<sub>tm</sub>, T1<sub>tm</sub>, and T2<sub>tm</sub> as appropriate to the phase of translation (see Fig. 1). In terms of T2<sub>tm</sub>, we address some of these approaches below.

As much as we cannot simply throw the science “over the wall” and expect the generation of appropriate interventions, nor can we simply throw these interventions “over the wall”. Whilst the models of translational research above have been elaborated beyond the dyadic T1 and T2 translational research, concern with the processes by which knowledge is translated into practice in the *real world*—that is, the messy, contingent, ambiguous, peopled world—have been taken up by a range of disciplines under the rubrics dissemination, implementation, and most recently, improvement science, i.e. T2<sub>tm</sub>. The important distinction between research translation and these variants of implementation/improvement science is the focus of the former on the *content* of the science and *form* of the interventions generated and of the latter

on the *processes* of moving the resultant science and interventions across the research translation continuum into practice. Key amongst these approaches is *knowledge translation* (KT) which is defined by the Canadian Institute for Health Research (CIHR) as “a dynamic and iterative process that includes synthesis, dissemination, exchange and ethically sound application of knowledge to improve the health of Canadians, provide more effective health services and products and strengthen the health care system” (CIHR 2011). With adjustments for citizenship this is the definition used in the international literature and in the Canadian context goes under the banner ‘better research, better decisions, better health’. KT has an inherently democratising orientation. KT draws upon a long history within health promotion, health education and community development of engaging communities, health care providers and policy makers with health care issues, action and change. These multiple perspectives share a common understanding of the translation of knowledge to practice as a complex dynamic social process. Moreover, these are *public*, not simply *scientific* issues. As with Nowotny et al.’s (2001, 2006) call for transdisciplinary knowledge generation, the forums for discussion and development of the translation of biobanking knowledge necessarily involve the representation and participation of scientists with industry, government and health service stakeholders as well as research participants and the public. Being neither the expert forums of academic conferences, select committees and health department committee meetings nor the public forums of patient groups, public meetings and demonstrations, these are, instead, what Callon et al. (2009) describe as *hybrid forums*. Translation of biobanking science demands the active involvement and intersection of perspectives of the full range of stakeholders: that is, the co-evolution of translational research.

Knowledge translation under the CIHR definition “takes place within a complex system of interactions between researchers and knowledge users which may vary in intensity, complexity and level of engagement depending on the nature of the research and the findings as well as the needs of the particular knowledge user” (CIHR 2011). KT offers a number of strategies of translation and exchange, the primary mode of which is a knowledge-to-action framework, which incorporate cycles of development that iterate between problem identification, knowledge synthesis, contextualising, tailoring and adapting knowledge, product and tools development, and evaluating and sustaining knowledge use (Graham and Tetroe 2007; Kitson and Straus 2009). Practices of KT include: use of consensus conferences or expert panels, systematic reviews, narrative syntheses, meta-analyses, meta-syntheses and practice guidelines to contextualise and integrate research findings; dissemination via briefings and educational sessions with stakeholders, including patients, practitioners and policy

makers, engaging knowledge users in developing and executing dissemination/implementation plans, and media engagement; knowledge exchange through interaction between knowledge users and researchers and results in mutual learning through the process of planning, producing, disseminating, and applying existing or new research in decision-making; and, are consistent with ethical principles and norms, social values, as well as legal and other regulatory frameworks (Kitson and Straus 2009). Empirical studies of KT strategies demonstrate their promise, but also their challenges (Löwy 1996; Mitton et al. 2009; Mitton et al. 2007) and such networks demand the reshaping of some members’ normative cultures and beliefs. In one example, the processes were inclusive and facilitated the involvement of non-academic actors, but tended to conceal political tensions (Lehoux et al. 2008). The resultant science was a pragmatic one whose intention was not to provide a critique or to produce unconventional knowledge (Lehoux et al. 2008); a *civilised* science (Lehoux et al. 2008). Although we must guard against an untrammelled science, an overly civilised science runs the risk of quashing innovation. Limits may impede bioscientific creativity and imagination but the challenges offered by transdisciplinarity can become a driving force for creativity (Nowotny 2008). As in biobanking science, the challenges of ethico-legal scholars (Cambon-Thomsen et al. 2007; Gottweis and Lauss 2010; Hoeyer 2008; Knoppers and Chadwick 2005) related to privacy and security of data have led to development of new technologies (Wolfson et al. 2010).

Studies of implementation suggest a range of challenges to moving new bioscience knowledge into practice. Kitson and Straus (2009) demonstrate inherent resistance in many organisations to embrace change and that most people (including professionals) operate most of the time on ‘automatic pilot’ and will unconsciously adapt to worsening conditions or tolerate diminishing standards. Denis et al. (2007) identifies the organisational components as necessary for KT and implementation—knowledge as: capabilities (support to use and diffuse knowledge, e.g. knowledge brokers, training); process (e.g. knowledge networks, communities of practice, integration of KT in research); and codification (e.g. practice guidelines, quality indicators, performance management systems). Normative Process Theory (May 2006; May et al. 2009) suggests that interventions only become routinely embedded (implemented and integrated) in their social, organisational and professional contexts as the result of people working, individually and collectively, to do so. Successfully achieving this requires an understanding of: how people make sense of the action(s) to be implemented; their attitudes; and their intentions. Importantly, implementation requires “continuous investment by people in ensembles of action that carry forward in time and space” (May et al. 2009).



Research translation is a contextual and cumulative process: one that is necessarily dynamic and interactive; one that involves multiple actors. Research translation is broadly conceived as one in which there is a contemporaneous “exploration of new avenues of research occurring in the same contexts in which new interventions are being developed and applications anticipated” (O’Malley and Stotz 2011). Interventions are ideally part of an iterative cycle of translation, incorporating further intervention and reintegration. Implementation/improvement science is predominantly focused on the T2 (i.e. T2–T4) phases of the translational research continuum. Arguably, in biobanking science, the methods of KT as described here are necessary across the continuum to include  $T0_{tm}$  and  $T1_{tm}$  methodologies to complete and maintain the cycle of translation. What is not yet entirely clear is just how to operationalise the translation cycle. Certainly knowledge generation will be collaborative, interactive and transdisciplinary. Practical methodologies of translation (including KT) in biobanking will likely include collaborative multiple stakeholder networks (*hybrid forums*) to build on the existing transdisciplinary science networks.

Addressing the complexity of bioscience knowledge generation and its implementation is crucial for robust understandings of translational research. This is arguably more so for biobanking science. Whilst biobanks undoubtedly provide a fundamental resource for both clinical and public health practice their potentiating ontology—that their outputs are perpetually a *promise* of scientific knowledge generation (Borup et al. 2006; Brown and Michael 2003; Martin et al. 2008)—renders translation rather less straightforward than drug discovery and treatment implementation. Biobanking science therefore provides a perfect counterpoint against which to test the bounds of thinking in terms of knowledge generation and its translation. We must understand translation more expansively to envisage how translational practice might be broadened. The beginnings of a robust model are to be found across the range of considerations of translational research and knowledge generation described above, both from within bioscience and without.

### Conclusion: translational research and the promise of biobanking science

In this paper, we described the distinctiveness of population biobanking and defined distinctions between a range of disciplinary approaches to translation in/of research. That we must imagine instead, for translational research for biobanks, a new paradigm: one that takes us from bench to bedside to backyard and beyond; that is attentive to the social and political context of translational science; and is cognisant of *all* the players in that process be they research-

ers, health professionals, policy makers, industry representatives or members of the public, amongst others. Although the question of just how we achieve this remains open, it is essential to direct our attention to these critical issues in the current phase of biobank development. We have outlined some key components here of an approach that requires further development. We need to explore how we can ensure that the maturing biobanks platforms are appropriately set up to optimise their ultimate value as drivers of translational change. There is a vital need to embed biobanks within health systems and to ensure involvement of policy makers, and health care providers early in the process. These developments will only succeed through cooperation. The depth and breadth of information that is ultimately required is so extensive that we will *never* have as much as we might ideally like. Consequently, it is to everybody’s advantage that biobank builders freely share their knowledge and expertise. The long-term management and use of data are social processes with ethico-legal implications. Much of the management of the science must focus on these issues rather than on the detailed bioscience itself. This demands the development and provision of new methods and tools, and the active co-involvement of bioscientists, clinical scientists, population and public health scientists, social scientists, and experts in information management and technology, and ethico-legal issues. To optimise ultimate outcomes, it should also involve the general public, health care providers and strategists, politicians and industry. No single group can possibly do the work that is required, and no single group can truly dominate or lead. Optimising the translation of biobank science into societal benefit is necessarily a *transdisciplinary* project.

**Acknowledgments** The work supporting the preparation of this manuscript is a core element of the research programme of P<sup>3</sup>G (the Public Population Project in Genomics) funded by Genome Canada and Genome Quebec. The methodological research programme at the University of Leicester focusing on Biobanking Science and the Sociology of Biobanking is funded jointly under the BioSHaRE-EU project (European Commission, FP7, #261433), Wellcome Trust Supplementary Grant #086160/Z/08/A, and joint MRC/Wellcome Trust Project Grant #G1001799/#WT095219MA.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

### References

- Antoine FS (1991) 1971–1991: biological therapy moves from bench to bedside. *J Natl Cancer Inst* 83:530
- Armstrong R, Waters E, Roberts H, Oliver S, Popay J (2006) The role and theoretical evolution of knowledge translation and exchange in public health. *J Public Health* 28:384

- BBMRI (2010) Biobanking and biomolecular research resources infrastructure. <http://www.bbMRI.eu/>
- BioSHARE-EU (2011) Biobank standardisation and harmonisation for research excellence in the European Union European Union—FP7 Program
- Borugian MJ, Robson P, Fortier I, Parker L, McLaughlin J, Knoppers BM, Bedard K, Gallagher RP, Sinclair S, Ferretti V, Whelan H, Hoskin D, Potter JD (2010) The Canadian partnership for tomorrow project: building a pan-Canadian research platform for disease prevention. *CMAJ* 182(11):1197–1201
- Borup M, Brown N, Konrad K, Van Lente H (2006) The sociology of expectations in science and technology. *Technol Anal Strategic Manage* 18(3/4):285–298
- Brown N, Michael M (2003) A sociology of expectations: retrospectively prospecting and prospecting retrospects. *Technol Anal Strategic Manage* 15:3–18
- Burke W, Burton H, Hall AE, Karmali M, Khoury MJ, Knoppers B, Meslin EM, Stanley F, Wright CF, Zimmern RL (2010) Extending the reach of public health genomics: what should be the agenda for public health in an era of genome-based and “personalized” medicine? *Genet Med* 12:785
- Burton PR, Tobin MD, Hopper JL (2005) Key concepts in genetic epidemiology. *Lancet* 366:941–951
- Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, Todd JA, Donnelly P, Barrett JC, Davison D, Easton D, Evans DM, Leung HT, Marchini JL, Morris AP, Spencer CC, Tobin MD, Attwood AP, Boorman JP, Cant B, Everson U, Hussey JM, Jolley JD, Knight AS, Koch K, Meech E, Nutland S, Prowse CV, Stevens HE, Taylor NC, Walters GR, Walker NM, Watkins NA, Winzer T, Jones RW, McArdle WL, Ring SM, Strachan DP, Pembrey M, Breen G, St Clair D, Caesar S, Gordon-Smith K, Jones L, Fraser C, Green EK, Grozeva D, Hamshere ML, Holmans PA, Jones IR, Kirov G, Moskvina V, Nikolov I, O’Donovan MC, Owen MJ, Collier DA, Elkin A, Farmer A, Williamson R, McGuffin P, Young AH, Ferrier IN, Ball SG, Balmforth AJ, Barrett JH, Bishop TD, Iles MM, Maqbool A, Yuldasheva N, Hall AS, Braund PS, Dixon RJ, Mangino M, Stevens S, Thompson JR, Bredin F, Tremelling M, Parkes M, Drummond H, Lees CW, Nimmo ER, Satsangi J, Fisher SA, Forbes A, Lewis CM, Onnie CM, Prescott NJ, Sanderson J, Matthew CG, Barbour J, Mohiuddin MK, Todhunter CE, Mansfield JC, Ahmad T, Cummings FR, Jewell DP et al (2007) Association scan of 14, 500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 39:1329–1337
- Burton PR, Hansell AL, Fortier I, Manolio TA, Khoury MJ, Little J, Elliott P (2009) Size matters: just how big is BIG? Quantifying realistic sample size requirements for human genome epidemiology. *Int J Epidemiol* 38:263–273
- Burton PR, Fortier I, Knoppers BM (2010) The Global Emergence of Epidemiological Biobanks: opportunities and challenges. Building the evidence for using genetic information to improve health and prevent disease. In: Khoury MJ, Gwinn M, Bradley L, Little J, Higgins JP, Ioannidis JP (eds) *Human genome epidemiology*, 2nd edn. Oxford University Press, New York
- Callon M (1999) The role of lay people in the production and dissemination of scientific knowledge. *Sci Technol Soc* 4(1):81
- Callon M, Larédo P, Rabeharisoa V, Gonard T, Leray T (1992) The management and evaluation of technological programs and the dynamics of techno-economic networks: the case of the AFME. *Res Policy* 21:215–236
- Callon M, Lascoumes P, Barthe Y (2009) *Acting in an uncertain world: an essay on technical democracy*. MIT Press, Cambridge
- Cambon-Thomsen A, Rial-Sebbag E, Knoppers BM (2007) Trends in ethical and legal frameworks for the use of human biobanks. *Eur Respir J* 30:373
- Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P (2000) Framework for design and evaluation of complex interventions to improve health. *BMJ* 321:694
- Carter C (1977) Monogenic disorders. *J Med Genet* 14:316–320
- CIHR (2011) Canadian Institute of Health Research, vol 2011
- Clayton D, McKeigue PM (2001) Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet* 358:1356–1360
- Collins FS (2004) The case for a US prospective cohort study of genes and environment. *Nature* 429:475–477
- Cooksey DS, Britain G, Treasury H (2006) A review of UK health research funding. Stationery Office Books (TSO)
- Cordell HJ, Clayton DG (2005) Genetic association studies. *Lancet* 366:1121–1131
- Cox D (2011) BBMRI: the industry perspective. <http://www.medical-device-network.com/features/feature97507/>. vol 2011
- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M (2008) Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 337:a1655
- Davey Smith G (2006) Randomised by (your) god: robust inference from an observational study design. *J Epidemiol Commun Health* 60:382–388
- Davey Smith G, Ebrahim S (2003) ‘Mendelian randomization’: can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 32:1–22
- Davey Smith G, Ebrahim S, Lewis S, Hansell AL, Palmer LJ, Burton PR (2005) Genetic epidemiology and public health: hope, hype, and future prospects. *Lancet* 366:1484–1498
- Demir I (2008) Incommensurabilities in the work of Thomas Kuhn. *Stud Hist Philos Sci Part A* 39:133–142
- Demir I (2011) Lost in translation? Try second language learning: understanding movements of ideas and practices across time and space. *J Hist Sociol* 24:9–26
- Denis JL, Langley A, Rouleau L (2007) Strategizing in pluralistic contexts: rethinking theoretical frames. *Human Relat* 60:179
- Devlin B, Roeder K, Wasserman L (2001) Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol* 60:155–166
- Didelez V, Sheehan N (2007) Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res* 16:309–330
- Dixon-Woods BC, Aveling EL, Goeschel CA, Pronovost PJ (2001) Explaining Michigan: developing an ex post theory of a quality improvement program. *Milbank Q* 89:167–205
- Doll R, Hill A (1952) A study of the aetiology of carcinoma of the lung. *BMJ* 2:1271–1286
- Doll R, Hill A (1964) Mortality in relation to smoking: 10 years’ observation of British doctors. *BMJ* 1(1399–1410):1460–1467
- Easton DF, Pooley KA, Dunning AM, Pharoah PDP, Thompson D, Ballinger DG, Struwing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odehrey F, Shen C-Y, Wu P-E, Wang H-C, Eccles D, Evans DG, Peto J, Fletcher O, Johnson N, Seal S, Stratton MR, Rahman N, Chenevix-Trench G, Bojesen SE, Nordestgaard BG, Axelsson CK, Garcia-Closas M, Brinton L, Chanock S, Lissowska J, Peplonska B, Nevanlinna H, Fagerholm R, Eerola H, Kang D, Yoo K-Y, Noh D-Y, Ahn S-H, Hunter DJ, Hankinson SE, Cox DG, Hall P, Wedren S, Liu J, Low Y-L, Bogdanova N, Schurmann P, Dork T, Tollenaar RAEM, Jacobi CE, Devilee P, Klijn JGM, Sigurdson AJ, Doody MM, Alexander BH, Zhang J, Cox A, Brock IW, MacPherson G, Reed MWR, Couch FJ, Goode EL, Olson JE, Meijers-Heijboer H, van den Ouweland A, Uitterlinden A, Rivadeneira F, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Hopper JL, McCredie M, Southey M, Giles GG, Schroen C, Justenhoven C, Brauch H, Hamann U,

- Ko Y-D, Spurdle AB, Beesley J, Chen X, Mannermaa A, Kosma V-M, Kataja V, Hartikainen J, Day NE et al (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 447:1087–1093
- Ebrahim S, Montaner D, Lawlor DA (2004) Clustering of risk factors and social class in childhood and adulthood in British women's heart and health study: cross sectional analysis. *Bmj* 328:861
- European Strategy Forum on Research Infrastructures (2006) European Roadmap for Research Infrastructures—Report 2006. European Communities, Brussels
- Farnworth A, Robson SC, Thomson RG, Watson DB, Murtagh MJ (2008) Decision support for women choosing mode of delivery after a previous caesarean section: a developmental study. *Patient Educ Couns* 71:116–124
- Fortier I, Burton PR, Robson PJ, Ferretti V, Little J, L'Heureux F, Deschenes M, Knoppers BM, Doiron D, Keers JC, Linksted P, Harris JR, Lachance G, Boileau C, Pedersen NL, Hamilton CM, Hveem K, Borugian MJ, Gallagher RP, McLaughlin J, Parker L, Potter JD, Gallacher J, Kaaks R, Liu B, Sprosen T, Vilain A, Atkinson SA, Rengifo A, Morton R, Metspalu A, Wichmann HE, Tremblay M, Chisholm RL, Garcia-Montero A, Hillege H, Litton JE, Palmer LJ, Perola M, Wolfenbuttel BH, Peltonen L, Hudson TJ (2010) Quality, quantity and harmony: the DataSHaPER approach to integrating data across bioclinical studies. *Int J Epidemiol* 39:1383–1393
- Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JRB, Elliott KS, Lango H, Rayner NW, Shields B, Harries LW, Barrett JC, Ellard S, Groves CJ, Knight B, Patch A-M, Ness AR, Ebrahim S, Lawlor DA, Ring SM, Ben-Shlomo Y, Jarvelin M-R, Sovio U, Bennett AJ, Melzer D, Ferrucci L, Loos RJF, Barroso I, Wareham NJ, Karpe F, Owen KR, Cardon LR, Walker M, Hitman GA, Palmer CNA, Doney ASF, Morris AD, Davey-Smith G, Consortium TWTC, Hattersley AT, McCarthy MI (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316(5826):889–894
- Geels FW (2005) Processes and patterns in transitions and system innovations: refining the co-evolutionary multi-level perspective. *Technol Forecast Social Change* 72:681–696
- Gen2Phen (2010) Knowledge centre. <http://www.gen2phen.org/>
- Gottweis H, Lauss G (2010) Biobank governance in the post-genomic age. *Personalized Med* 7:187–195
- Graham ID, Tetroe J (2007) Some theoretical underpinnings of knowledge translation. *Acad Emerg Med* 14:936–941
- Green ED, Guyer MS (2011) Charting a course for genomic medicine from base pairs to bedside. *Nature* 470:204–213
- Hammersley M, Atkinson P (2007) *Ethnography: principles in practice*. Taylor & Francis, London
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 106:9362–9367
- Hoeyer K (2008) The ethics of research biobanking: a critical review of the literature. *Biotechnol Genet Eng Rev* 25:429–452
- ISBER (2005) Best practices for repositories I: collection, storage, and retrieval of human biological materials for research. *Cell Preserv Technol* 3:5
- Jasanoff S (2004) *States of knowledge: the co-production of science and social order*. Routledge, London
- Khoury MJ (2004) The case for a global human genome epidemiology initiative. *Nat Genet* 36:1027–1028
- Khoury M (2010) Dealing with the evidence dilemma in genomics and personalized medicine. *Clin Pharmacol Ther* 87:635–638
- Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L (2007) The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genet Med* 9:665
- Khoury MJ, Feero WG, Reyes M, Citrin T, Freedman A, Leonard D (2009a) The genomic applications in practice and prevention network. *Genet Med* 11:488
- Khoury MJ, Gwinn M, Bradley L, Little J, Higgins JP, Ioannidis JP (2009b) *Human genome epidemiology*, 2nd edn. Oxford University Press, New York
- Khoury MJ, Bowen MS, Burke W, Coates RJ, Dowling NF, Evans JP, Reyes M, St Pierre J (2011) Current priorities for public health practice in addressing the role of human genomics in improving population health. *Am J Prev Med* 40:486–493
- Kirch W (2008) *Encyclopaedia of public health*. Springer Science and Business Media, New York
- Kitson AS, Straus SE (2009) *Knowledge translation in health care: moving from evidence to practice*. Wiley, Blackwell
- Knoppers BM, Chadwick R (2005) Human genetic research: emerging trends in ethics. *Nat Rev Genet* 6:75–79
- Knoppers BM, Kent A (2006) Policy barriers in coherent population-based research. *Nat Rev Genet* 7: 8–8
- Knoppers BM, Fortier I, Legault D, Burton P (2008) The Public Population Project in Genomics (P3G): a proof of concept? *Eur J Hum Genet* 16:664–665
- Lehoux P, Daudelin G, Denis JL, Miller F (2008) Scientists and policy-makers at work: listening to epistemic conversations in a genetics science network. *Sci Public Policy* 35:207–220
- Litton JE, Muilu J, Bjorklund A, Leinonen A, Pedersen NL (2003) Data modeling and data communication in GenomEUtwin. *Twin Res* 6:383–390
- Löwy I (1996) *Between bench and bedside: science, healing, and interleukin-2 in a cancer ward*. Harvard Univ Press, Harvard
- Manolio TA, Bailey-Wilson JE, Collins FS (2006) Genes, environment and the value of prospective cohort studies. *Nat Rev Genet* 7:812–820
- Manolio TA, Brooks LD, Collins FS (2008) A HapMap harvest of insights into the genetics of common disease. *J Clin Invest* 118:1590–1605
- Martin P, Brown N, Kraft A (2008) From bedside to bench? Communities of promise, translational research and the making of blood stem cells. *Sci Cult* 17:29–41
- May C (2006) A rational model for assessing and evaluating complex interventions in health care. *BMC Health Services Res* 6:86
- May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, Rapley T, Ballini L, Ong BN, Rogers A (2009) Development of a theory of implementation and integration: normalization process theory. *Implement Sci* 4:29
- Merikangas KR, Risch N (2003) Genomic priorities and public health. *Science* 302:599–601
- Merz B (1985) Nobelists take genetics from bench to bedside. *JAMA* 254:3161
- Mitton C, Adair CE, McKenzie E, Patten SB, Perry BW (2007) Knowledge transfer and exchange: review and synthesis of the literature. *Milbank Q* 85:729–768
- Mitton C, Adair CE, McKenzie E, Patten S, Wayne-Perry B, Smith N (2009) Designing a knowledge transfer and exchange strategy for the Alberta Depression Initiative: contributions of qualitative research with key stakeholders. *Int J Ment Health Syst* 3:1–10
- Moore J, Asselbergs F, Williams S (2010) Bioinformatics challenges for genome-wide association studies. *Bioinformatics* 26:445–455
- Moore HM, Compton C, Alper J, Vaught JB (2011) International approaches to advancing biospecimen science. *Cancer Epidemiol Biomarkers Prev* 20(5):729–732
- Moreira T (2009) Hope and truth in drug development and evaluation in Alzheimer's Disease. In: Bellinger JF (ed) *Treating dementia: do we have a pill for it?* Johns Hopkins University Press, Baltimore, pp 210–230

- Moreira T, Palladino P (2005) Between truth and hope: on Parkinson's disease, neurotransplantation and the production of the 'self'. *Hist Human Sci* 18:55
- Murtagh MJ, Hepworth J (1997) Implications of menopause for public health. In: HERDSA (ed) National Rural Public Health Forum. HERDSA, Canberra
- Murtagh MJ, Hepworth J (2003) Menopause as a long term risk to health: implications of general practitioner accounts of prevention for women's choice and decision making. *Social Health Illn* 25:185–207
- Murtagh MJ, Thomson RG, May CR, Rapley T, Heaven BR, Graham RH, Kaner EF, Stobbart L, Eccles MP (2007) Qualitative methods in a randomised controlled trial: the role of an integrated qualitative process evaluation in providing evidence to discontinue the intervention in one arm of a trial of a decision support tool. *Qual Saf Health Care* 16:224
- Newton-Cheh C, Eijgelsheim M, Rice KM, de Bakker PI, Yin X, Estrada K, Bis JC, Marciante K, Rivadeneira F, Noseworthy PA, Sotoodehnia N, Smith NL, Rotter JJ, Kors JA, Witteman JC, Hofman A, Heckbert SR, O'Donnell CJ, Uitterlinden AG, Psaty BM, Lumley T, Larson MG, Stricker BH (2009a) Common variants at ten loci influence QT interval duration in the QTGEN Study. *Nat Genet* 41:399–406
- Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, Najjar SS, Zhao JH, Heath SC, Eyheramendy S, Papadakis K, Voight BF, Scott LJ, Zhang F, Farrall M, Tanaka T, Wallace C, Chambers JC, Khaw KT, Nilsson P, van der Harst P, Polidoro S, Grobbee DE, Onland-Moret NC, Bots ML, Wain LV, Elliott KS, Teumer A, Luan J, Lucas G, Kuusisto J, Burton PR, Hadley D, McArdle WL, Brown M, Dominiczak A, Newhouse SJ, Samani NJ, Webster J, Zeggini E, Beckmann JS, Bergmann S, Lim N, Song K, Vollenweider P, Waeber G, Waterworth DM, Yuan X, Groop L, Orho-Melander M, Allione A, Di Gregorio A, Guarrera S, Panico S, Ricceri F, Romanazzi V, Sacerdote C, Vineis P, Barroso I, Sandhu MS, Luben RN, Crawford GJ, Jousilahti P, Perola M, Boehnke M, Bonnycastle LL, Collins FS, Jackson AU, Mohlke KL, Stringham HM, Valle TT, Willer CJ, Bergman RN, Morken MA, Doring A, Gieger C, Illig T, Meitinger T, Org E, Pfeufer A, Wichmann HE, Kathiresan S, Marrugat J, O'Donnell CJ, Schwartz SM, Siscovick DS, Subirana I, Freimer NB, Hartikainen AL, McCarthy MI, O'Reilly PF, Peltonen L, Pouta A, de Jong PE, Snieder H, van Gilst WH, Clarke R, Goel A, Hamsten A, Peden JF et al (2009b) Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet* 41:666–676
- NIH (2011) <http://www.grants.nih.gov/grants/glossary.htm#T>
- Nowotny H (2008) *Insatiable curiosity: innovation in a fragile future*. MIT Press, Cambridge
- Nowotny H, Scott P, Gibbons M (2001) *Re-thinking science: knowledge and the public in an age of uncertainty*. Polity Press, Cambridge
- Nowotny H, Scott P, Gibbons M (2006) *Re-thinking science. Knowledge creation, diffusion, and use in innovation networks and knowledge clusters: a comparative systems approach across the United States, Europe, and Asia*, vol 39
- O'Malley MA, Stotz K (2011) Intervention, integration and translation in obesity research: Genetic, developmental and metaorganismal approaches. *Philos Ethics Humanities Med* 6:2
- P3G Consortium, Church G, Heeney C, Hawkins N, de Vries J, Boddington P, Kaye J, Bobrow M, Weir B (2009) Public Access to Genome-Wide Data: five views on balancing research with privacy and protection. *PLoS Genet* 5:e1000665
- P3G Observatory (2009). <http://www.p3gobservatory.org>
- Palmer TM, Thompson JR, Tobin MD, Sheehan NA, Burton PR (2008) Adjusting for bias and unmeasured confounding in Mendelian randomization studies with binary responses. *Int J Epidemiol* 1161–1168
- Peakman TC, Elliott P (2008) The UK Biobank sample handling and storage validation studies. *Int J Epidemiol* 37(Suppl 1):i2–i6
- Pisani E, AbouZahr C (2010) Sharing health data: good intentions are not enough. *Bull World Health Organ* 88:462–466
- Pritchard JK, Cox NJ (2002) The allelic architecture of human disease genes: common disease-common variant or not? *Hum Mol Genet* 11:2417–2423
- Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, Zhao JH, Ramasamy A, Zhai G, Vitart V, Huffman JE, Igl W, Albrecht E, Deloukas P, Henderson J, Granell R, McArdle WL, Rudnicka AR, Barroso I, Loos RJ, Wareham NJ, Mustelin L, Rantanen T, Surakka I, Imboden M, Wichmann HE, Grkovic I, Jankovic S, Zgaga L, Hartikainen AL, Peltonen L, Gyllenstein U, Johansson A, Zaboli G, Campbell H, Wild SH, Wilson JF, Glaser S, Homuth G, Volzke H, Mangino M, Soranzo N, Spector TD, Polasek O, Rudan I, Wright AF, Heliouvaara M, Ripatti S, Pouta A, Nalwai AT, Olin AC, Toren K, Cooper MN, James AL, Palmer LJ, Hingorani AD, Wannamethee SG, Whincup PH, Smith GD, Ebrahim S, McKeever TM, Pavord ID, Macleod AK, Morris AD, Porteous DJ, Cooper C, Dennison E, Shaheen S, Karrasch S, Schnabel E, Schulz H, Grallert H, Bouatia-Naji N, Delplanque J, Froguel P, Blakey JD, Britton JR, Morris RW, Holloway JW, Lawlor DA, Hui J, Nyberg F, Jarvelin MR, Jackson C, Kahonen M, Kaprio J, Probst-Hensch NM, Koch B, Hayward C, Evans DM, Elliott P, Strachan DP, Hall IP, Tobin MD (2009) Genome-wide association study identifies five loci associated with lung function. *Nat Genet* 42:36–44
- Saxena R, Voight BF, Lyssenko V, Burt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskiran MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316:1331–1336
- Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316:1341–1345
- Sheehan NA, Didelez V, Burton PR, Tobin MD (2008) Mendelian randomisation and causal inference in observational epidemiology. *PLoS Med* 5:e177
- Spencer CC, Su Z, Donnelly P, Marchini J (2009) Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet* 5:e1000477
- Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, Albers-Akkers MT, Swinkels DW, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Godino J, Garcia-Prats MD, Polo E, Tres A, Mouy M, Saemundsdottir J, Backman VM, Gudmundsson L, Kristjansson K, Bergthorsson JT, Kostic J, Frigge ML, Geller F, Gudbjartsson D, Sigurdsson H, Jonsdottir T, Hrafnkelsson J, Johannsson J,

- Sveinsson T, Myrdal G, Grimsson HN, Jonsson T, von Holst S, Werelius B, Margolin S, Lindblom A, Mayordomo JI, Haiman CA, Kiemeny LA, Johannsson OT, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K (2007) Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 39:865–869
- Stover PJ, Harlan WR, Hammond JA, Hendershot T, Hamilton CM (2010) PhenX: a toolkit for interdisciplinary genetics research. *Curr Opin Lipidol* 21(2):136–140
- Taubes G (1995) Epidemiology faces its limits. *Science* 269:164–169
- Thomson R, Murtagh M, Khaw FM (2005) Tensions in public health policy: patient engagement, evidence-based public health and health inequalities. *Qual Saf Health Care* 14:398
- Timpson NJ, Lawlor DA, Harbord RM, Gaunt TR, Day IN, Palmer LJ, Hattersley AT, Ebrahim S, Lowe GD, Rumley A, Davey Smith G (2005) C-reactive protein and its role in metabolic syndrome: Mendelian randomisation study. *Lancet* 366:1954–1959
- Wallace S, Lazor S, Knoppers BM (2009) Consent and population genomics: the creation of generic tools. *IRB Ethics Human Res* 31:15–20
- Walport M, Brest P (2011) Sharing research data to improve public health. *Lancet* 377:537–539
- Wellcome\_Trust\_Case\_Control\_Consortium (2007) Genome-wide association study of 14, 000 cases of seven common diseases and 3, 000 shared controls. *Nature* 447:661–678
- Westfall JM, Mold J, Fagnan L (2007) Practice-based research—“Blue Highways” on the NIH roadmap. *JAMA* 297:403
- Wichmann HE et al (2011) Comprehensive catalogue of European biobanks. *Nat Biotechnol* (in press)
- Wolfson M, Wallace SE, Masca N, Rowe G, Sheehan NA, Ferretti V, Laflamme P, Tobin MD, Macleod J, Little J, Fortier I, Knoppers BM, Burton PR (2010) DataSHIELD: resolving a conflict in contemporary bioscience—performing a pooled analysis of individual-level data without sharing the data. *Int J Epidemiol* 39:1372–1382
- Wong MY, Day NE, Luan JA, Chan KP, Wareham NJ (2003) The detection of gene-environment interaction for continuous traits: should we deal with measurement error by bigger studies or better measurement? *Int J Epidemiol* 32:51–57
- Woolf SH (2008) The meaning of translational research and why it matters. *JAMA* 299:211
- World Health Organisation (2008) WHO Report on the global tobacco epidemic. World Health Organisation, Geneva
- Yuille M, van Ommen GJ, Brechot C, Cambon-Thomsen A, Dagher G, Landegren U, Litton JE, Pasterk M, Peltonen L, Taussig M, Wichmann HE, Zatloukal K (2008) Biobanking for Europe. *Brief Bioinform* 9:14–24
- Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JRB, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney ASF, WTCCC, McCarthy MI, Hattersley AT (2007) Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316:1336–1339
- Zerhouni E (2003) Medicine. The NIH Roadmap. *Science* 302:63–72
- Zerhouni EA (2005) Translational and clinical science—time for a new vision. *N Engl J Med* 353:1621–1623
- Zimmern RL, Brice PC (2009) Realizing the potential of genomics: translation is not translational research. *Genet Med* 11:898
- Zondervan KT, Cardon LR (2007) Designing candidate gene and genome-wide case-control association studies. *Nat Protoc* 2:2492–2501