Review

CellPress

Genomics and personalised whole-of-life healthcare

Denis C. Bauer[1], Clara Gaff[2], Marcel E. Dinger[3], Melody Caramins[4,5], Fabian A. Buske[6], Michael Fenech[7,8], David Hansen[9], and Lynne Cobiac[7]

[1] CSIRO Preventative Health Flagship and CSIRO Computational Informatics, Sydney, NSW 2113, Australia
[2] The Walter and Eliza Hall Institute, Parkville, VIC 3052, Australia
[3] Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Australia
[4] School of Medical Sciences, University of New South Wales, Sydney, Australia
[5] Laverty/SDS Pathology, 60 Waterloo Road, North Ryde, NSW 2113, Australia
[6] Cancer Research Division, Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Australia
[7] CSIRO Preventative Health National Research Flagship, Adelaide, SA 5000, Australia
[8] CSIRO Animal, Food, and Health Sciences, Gate 13 Kintore Avenue, Adelaide, SA 5000, Australia
[9] Australian E-Health Research Centre and CSIRO Computational Informatics, Herston, QLD 4029, Australia

Genome sequencing has the potential for stratified cancer treatment and improved diagnostics for rare disorders. However, sequencing needs to be utilised in risk stratification on a population scale to deepen the impact on the health system by addressing common diseases, where individual genomic variants have variable penetrance and minor impact. As the accuracy of genomic risk predictors is bounded by heritability, environmental factors such as diet, lifestyle, and microbiome have to be considered. Large-scale, longitudinal research programmes need to study the intrinsic properties between both genetics and environment to unravel their risk contribution. During this discovery process, frameworks need to be established to counteract unrealistic expectations. Sufficient scientific evidence is needed to interpret sources of uncertainty and inform decision making for clinical management and personal health.

Technological advances propel personalised healthcare

Providing tailored healthcare to optimally cater for the specific needs of an individual lies at the heart of medical practice. Today, unprecedented computational capabilities and high-throughput data collection methods promise a new era of personalised, evidence-based healthcare, utilising individual genetic (see Glossary) or genomic testing to tailor health management. The technological potential has been demonstrated, among others, for genome-informed treatment [1], prenatal diagnoses [2], and research relating lifestyle choices and cancer risk [3]. These successes have inspired stakeholders with commercial interest to provide direct-to-consumer (DTC) testing for genetic variants associated with certain health conditions.

A recent survey in North America estimates the general interest in DTC genetic testing to be 51–80% [4]. Another study assessing psychological impact of information obtained from DTC genetic testing found that there was no significant change in anxiety up to 12 months after testing, even in the absence of professional counselling [5]. However, there was also no significant change in diet, lifestyle, or screening behaviour among those taking the tests, unless the participants shared their result with their physician. Those that did so (one-third of the participants) generally had a higher genetic risk for the 28 conditions tested by Navigenics and were subsequently more likely to complete screening tests [5].

As forecasted by Knoppers *et al.* [6], the growing yet insufficiently regulated commercial market has let stakeholders overstate the potential benefits of genetic testing

Glossary

Electronic Health Driven Genome Research (EDGR): utilizing EHRs for the cost-effective recruitment of a precise and well-annotated patient cohort.
Environment-wide association study (EWAS): a study designed to determine the association between a specific phenotype and environmental factors.
Exposome: combination of physical, chemical, nutritional, lifestyle, and psychosocial exposure factors.
Genetic diagnosis: determining the genotype of a known disease marker.
Genetic discrimination: different treatment based on a person's actual or assumed genetic makeup.
Genetic screening: obtaining the genomic profile of an individual.
Genome-wide association study (GWAS): a study designed to determine the association between a specific phenotype and the genotype of common genetic variants.
Genomic profile: set of genetic variations observed in an individual.
Hospital electronic health records (EHRs): medical data and treatment history in a machine-readable format shared with all relevant healthcare providers of the patient.
Personalised nutrition: adapting food and nutritional intake depending on the individual's genomic profile, life stage, and lifestyle to prevent disease or delay its onset.
Phenome-wide association study (PheWAS): a study designed to determine the association between specific genomic regions and the observed phenotypes.
Prevalence: proportion of a population found to have a specific disease condition.
Receiver operator curve (ROC): visualisation of sensitivity and specificity of a prediction method.
Risk stratification: subdividing a population into groups based on genetic profiles.

Corresponding author: Bauer, D.C. (Denis.Bauer@CSIRO.au).
Keywords: personalised health; preventative health; genomic profiling; clinical genomics; risk prediction; digital health services.

1471-4914/
 http://dx.doi.org/10.1016/j.molmed.2014.04.001

without adequate heed to the ethical and social implications. Simultaneously, policymakers have been put under substantial pressure to extend public screening programmes before the scientific merit of such tests has been adequately demonstrated [7]. This prompted the European Academies Science Advisory Council (EASAC) to submit a report to the European Union advising regulation of DTC genetic testing to exclude diagnostic, presymptomatic, and prenatal testing. It recommended that policymakers consider the full implications of introducing genomic sequencing in screening programmes [8]. In the USA, the FDA has stopped marketing campaigns featuring genetic tests until claims of their clinical utility are verified and cleared by the FDA (document number: GEN1300666).

However, this widens the gap between what is technologically possible and the current medical practice [9,10]. In light of this, we review the state-of-the-art in personalised healthcare, investigate the continuation to personalised whole-of-life health approaches, discuss technical issues that may prevent future realisation of these ideas, and outline how life science in conjunction with technological advances can improve clinical practice and enable people to 'self-manage in the face of physical, emotional, and social challenges' [11].

Genomic profiles improve personal diagnoses and risk prediction

Traditional testing for specific genetic aberrations has been estimated to lead to a diagnosis in 42% of cases attending a genetic clinic (8% could be diagnosed to have no genetic disorder), with the majority of diagnoses made at the time of the first clinician consultation [12]. For the remaining 50% who could not be diagnosed with the first test, the cost per diagnosis as the result of sequential traditional testing has been conservatively estimated at US $25 000 [12]. A clinically and economically beneficial alternative for patients with an unknown high penetrance genetic disorder (Figure 1A) is whole genome/exome sequencing as it can identify germline variants from across the genome simultaneously. Genomic sequencing can lead to a successful diagnosis in up to 50% of cases where traditional genetic testing failed [12]. This can provide a more precise diagnosis in individuals affected by a condition, which may inform new treatment options. The potential of this technology is only beginning to be realised with the majority of genomic variants as yet clinically uncharacterised, and therefore of unknown significance [13]. Collaborative efforts across genomic laboratories, such as the Human Variome Project [14] and ClinVar [15], play an

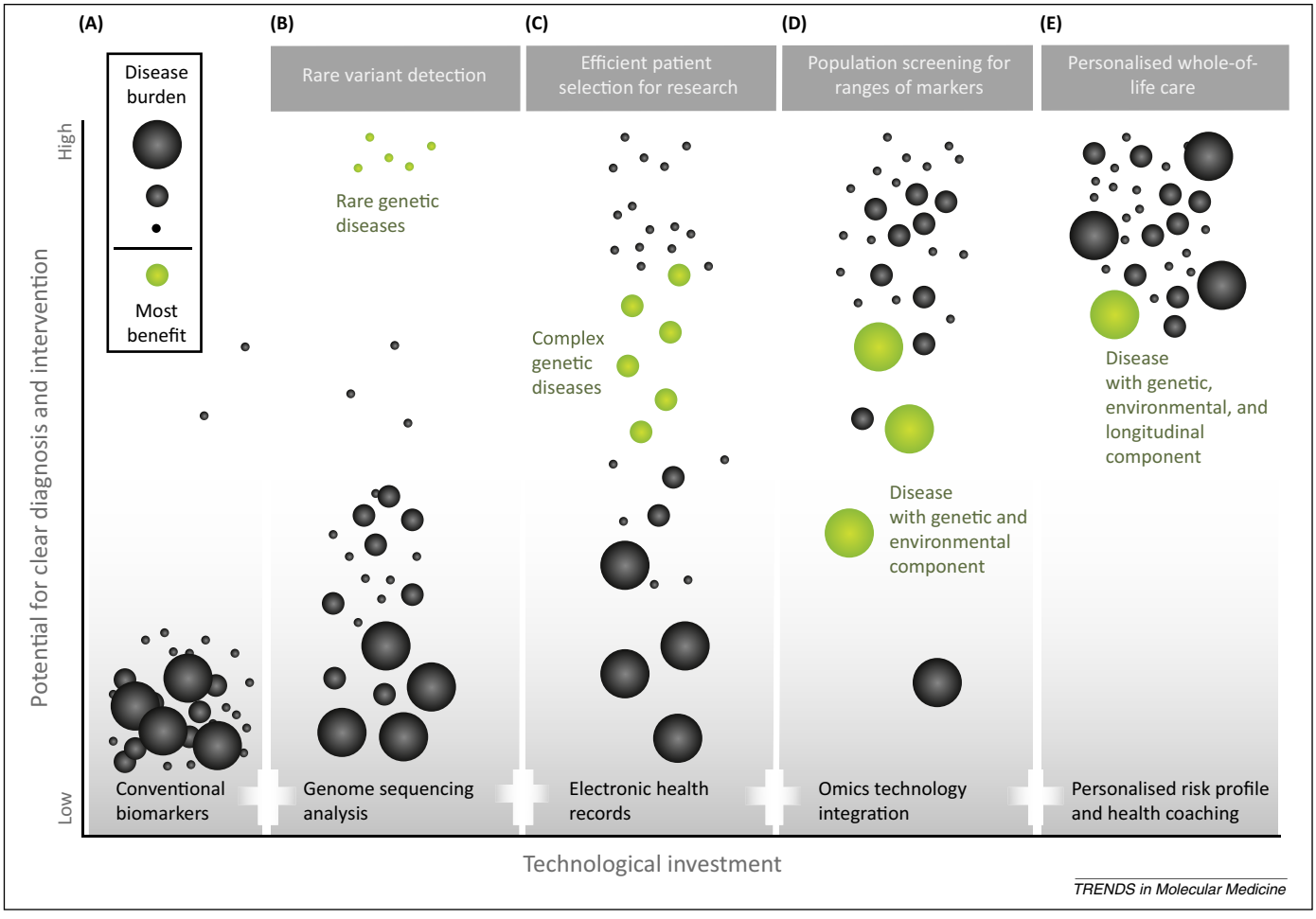

Figure 1. **(A–E)** Technological advancement towards whole-of-life care. Patients with rare genetic diseases are first to benefit from genomic sequencing-based testing as affected genes and pathways can be identified (B). Utilising electronic health records will improve diagnosis and treatment of complex genetic diseases, such as cancer, as informative patient groupings can be achieved more efficiently (C). Complex diseases with a clear environmental component, such as diabetes, will benefit from the integration of other omics-based assays, for example, DNA methylation (D); however, personalised preventative and intervention strategies (E) are required to manage the health of this high-burden multicausal group optimally.

important role in documenting genetic variants and their clinical associations. As these repositories of genotype–phenotype associations continue to expand, the diagnostic yield from genome sequencing will continue to grow.

Insights into cellular mechanisms, technological procedures, and clinical treatment from genetic testing are often directly applicable to somatic variants, such as those that can lead to the development of cancer. Current cancer treatments are largely determined by the organ affected; yet, genomic profiling shows enormous heterogeneity even among tumours with shared tissue origins [16]. As cancer is ultimately driven by genetic alterations, genomic tumour profiling is increasingly recognised as a valuable additional tool to inform treatment. A recent comprehensive survey of cancer genomes demonstrated a correlation between a patient's tumour genome and the clinical features of disease, drug response, and patient outcomes [16]. The collaborative efforts from The Cancer Genome Atlas (TCGA) has subsequently identified 291 driver genes acting on 3205 tumours from 12 different cancer types [17], which paves the way to driver gene-based tumour diagnosis and treatment advice in the future.

A third application of genomic testing is the prediction of future disease in healthy individuals. Currently, genotype-based disease risk prediction is limited to highly penetrant and relatively uncommon inherited diseases. For example, hereditary breast and ovarian cancer, caused by dominantly inherited mutations in the *BRCA1* and *BRCA2* genes, accounts for only 5% of breast cancer cases [18]. However, the majority of common conditions are complex multigene disorders arising from the interplay of multiple genetic variants. Each variant individually explains only a small proportion of the risk, but collectively and in concert with both identifiable and non-identifiable environmental and stochastic factors define the total risk. This is known as the 'liability threshold model' [19].

To provide informative testing for multiple variants for common complex disorders, large-scale studies are needed to assess the strength of the association between a specific combination of variants and the disease, as well as to determine the predictive value of the profile. However, making predictions from genomic profiles about an individual's complex disease risk to inform outcome or treatment will not be possible on currently feasible sample sizes [16,20] due to the 'curse of dimensionality' (Box 1). This is especially the case for complex diseases, as these are not determined by genetic profiles alone (discussed as 'missing heritability' in Box 2), requiring even more dimensions to be factored into the risk calculation [21].

Accordingly, more studies are needed to better understand the multifactorial contribution to disease of individual variants, interacting variants, and exposure risk factors before we can start to provide advice based on risk profiles. Efforts are currently underway with the initiation of the 'Human Phenome Project' [14] and the 'Evaluation of Genomic Applications in Practice and Prevention' (EGAPP) [22].

Box 1. Curse of dimensionality

In particular for inherited genetic disorders, meta-analyses of GWAS with genomic profiles of 10 000 samples demonstrate that the sample size would have to be orders of magnitude higher to accurately train a predictor for individual disease risks [60]. This is known as the 'curse of dimensionality' (the $P >> n$ problem), which states that prediction accuracy requires a training set to have a sufficiently large number of samples (individuals, n) to learn the contribution from each of the measured features (e.g., single nucleotide polymorphisms, P). Note, GWAS interrogate common variants and are hence, by design, powered to detect association with causal variants that are relatively common in the population [20] and therefore have little predictive value. Although sequencing studies can identify novel disease associations, the overall risk for common disorders explained by these low frequency variants is not expected to be major.

Box 2. Missing heritability

The heritability of many complex diseases has escaped explanation despite extensive genetic studies, primarily through GWAS. This phenomenon, dubbed as 'missing heritability' [61], has been debated to be the result of various potential mechanisms, including allelic architecture, rare variants, epistasis, parent of origin effects, and epigenetics (for reviews, see [62,63]). The latter is exemplified by twin studies showing variable disease penetrance in genetically identical individuals. The combination of environment and lifestyle choices has been shown to contribute at comparable levels to genetic loci and has been proposed to explain a large fraction of the observed variance that is not explained by GWAS hits [64].

Personal preventative strategies need to be based on risk stratification

Predicting an individual's disease risk is desirable for early intervention and preventive strategies. Assessing the individual risk is currently only feasible for rare diseases where biomarker or genetic tests are highly predictive. Some authors argue that technological progress will substantially increase sample sizes and computational ability, which will make risk prediction feasible for more common diseases where multiple loci contribute to the genetic burden [10].

However, to achieve higher precision of prediction, it will be necessary to incorporate the environmental exposure factors into risk calculations (Figure 1D). A recent environment-wide association study (EWAS) on type 2 diabetes mellitus (T2D) discovered a significant association with a wide range of exposure factors such as the pesticide-derivative heptachlor epoxide, polychlorinated biphenyls, and vitamins E and A [23]. Ashley *et al.* [24] extended this work to study adverse drug reactions. Despite difficulty in ascertaining causality, the potential for novel exposure factors of large effect associated with T2D justify the use of EWAS given that environmental factors can be found with effect sizes comparable to the best loci found by genome-wide association studies (GWAS) [25]. EWAS will become even more important once the capacity to define and quantify the complete exposome (i.e., the combination of physical, chemical, nutritional, lifestyle, and psychosocial exposure factors) becomes feasible [26–28].

Therefore, applying genetic testing as a screening rather than a diagnostic tool can identify people at high risk and focus preventative strategies, known as risk stratification. Take, for example, a genotype risk score for type 1 diabetes (T1D) based on a model using known loci. This score needs to be calibrated typically by means of a receiver operator curve (ROC) showing the proportion of cases detected versus the proportion of the population classified

as positive by the test. If the threshold is set so that it returns a positive result for 18% of the population, as T1D has a prevalence of 0.004, this means that for every 50 who receive preventative intervention one T1D case can be prevented, an improvement from originally one for every 250. In this way, a weak diagnostic test can be repurposed as a useful risk stratification tool [29].

Translating successful methodologies for cancer and rare genetic diseases to predict individual risk for common diseases remains challenging. However, utilising genetic profiles and whole-of-life exposure to stratify the population into risk groups will inform individual life-long preventative strategies.

Personalised nutrition needs to be underpinned with more evidence

Adapting food and nutritional intake to individual needs, depending on the individual's life stage, lifestyle, and life situation is desirable for disease prevention (Figure 1E) [30]. Although lifestyle interventions are effective for preventing T2D at any level of genetic risk, the benefit increases for individuals at high genetic susceptibility [31]. Hence, disease risk may be reduced by changing health behaviour in response to better understanding of all aspects that have an influence, such as an individual's genome, diet, and the human gut microbiota, which encompasses a complex ecosystem in the intestine with a profound impact on the host metabolism [30]. An attack on the entirety of this problem is likely to run directly into the 'curse of dimensionality'. It is hence not surprising that only modest progress has been made in this field [32]. Nevertheless, personalised nutrition has proven to be important and successful in the case of inherited or acquired metabolic disorders in which the molecular mechanisms are well established, such as in phenylketonuria, lactose intolerance, and coeliac disease [33–35].

Establishing robust dietary recommendations to prolong presymptomatic states is specifically complicated as emerging properties from this complex system can have unpredictable negative effects. For example, an increased intake of riboflavin may be recommended to lower blood pressure in those homozygous for the methylenetetrahydrofolate reductase (*MTHFR*) C677T polymorphism [36], yet increasing riboflavin in a low-folate physiological background can increase genomic instability and thus cancer risk [37]. Only long-term large-scale scientific studies can investigate which risk is higher and what intervention should be undertaken depending on genotype, lifestyle, or disease state (e.g., cancer survivor, smoker). In the interim, in the absence of reliable mathematical predictive models of nutrient–gene interactive effects of personalised intervention outcomes, it remains possible to utilise nutriome/nutrient array *in vitro* test systems to determine the optimal nutrient combinations to deliver a desired health outcome at the cellular level (e.g., genome stability, mitochondrial function) for an individual utilising their own cells (e.g., lymphocytes, fibroblasts, stem cells) [37].

Therefore, to tailor dietary advice around a person's genetic profile, well-designed studies interrogating specific hypotheses are necessary, as planned by the 'Micronutrients Genomics Project' [38].

Uptake of genomics

Genetic testing results for highly penetrant 'pathogenic' variants (mutations) have been used in clinical practice for several decades. These tests include preimplantation and prenatal diagnoses, paediatric and adult-onset genetic conditions, and predictive testing for highly penetrant adult-onset conditions and treatment of cancer [12]. Genetic testing has also been applied in screening programs of the general population (i.e., newborn screening) and of high-risk populations (e.g., Tay Sachs screening in the Jewish community). However, the application of high-throughput genomic technologies for these purposes is in its infancy. Currently, whole genome/exome sequencing is used for diagnoses, where other tests have failed to identify the causative mutation [39]. Recently, the National Institutes of Health (NIH) announced funding of research programmes to investigate the use of genomic sequencing in newborn screening presaging the use of these tests in population health programmes [40].

Realising the potential for better diagnosis and prevention will have challenges, however. Some of these challenges relate to the nature of genetic information, whereas others arise specifically from genomics.

Use of genomic sequencing by clinicians

The barriers to the use of genetic tests by healthcare practitioners have been reiterated in numerous studies [5,41,42]. Foremost among these is familiarity with genetic tests and their interpretation. For instance, a recent survey identified the uncertainty among clinicians about the available genetic tests, as well as lack of training and guidelines as the largest influences on the decision to include genetic testing into patient care when assessing predisposition to diseases and/or drug response [43]. Although it can be anticipated that this will be a barrier to genomic sequencing, the impact is likely to be compounded by characteristics specific to genomic technologies.

Traditionally, a patient sample is tested for specific somatic variants or heritable mutations in a specific gene (or set of genes). Identification of 'all' variants present in an individual requires sequencing of the genome or exome, followed by meticulous interrogation, curation, and clinical interpretation, with prioritisation of relevant variants with regard to the patient's clinical presentation. This task is complicated by the many variants found of potential but unknown significance, particularly if probands are considered in isolation. Patients who have most benefitted from genome sequencing have typically been subjects of gene identification studies attempting to identify causative variants that have not previously been described in association with the patient's clinical presentation [1]. Generally, these cases are examples of novel genetic disorders or due to atypical phenotypic presentations of known genetic disorders. Although this blurs the boundaries between clinical service and research, virtually all medical diagnoses are made with the support of laboratory evidence. The EGAPP initiative provides a framework for evaluating evidence for the transition of genomic tests from research to clinical practice [26]. Thereafter, clear guidelines for clinicians are necessary so that they can provide evidence-based advice to patients on therapeutic

choice or reduction of the risk attributable to genetic variants.

In single gene testing, the results generally relate only to the patient's condition or concern. Chromosomal microarrays were the first instance of a genomic test that was simultaneously a diagnostic and (potentially) predictive test. Genomic sequencing is an extension of this with potentially greater numbers of variants identified, some of which may be predictive of diseases or risks unrelated to the original clinical indication for which the test was performed. This introduces ethical dilemmas regarding the reporting of these 'incidental' findings and recommendations that known causative variants for 37 adult-onset, treatable diseases should always be reported in clinical testing have proven highly contentious [44].

Although preliminary data indicate that over 93% of patients choose to receive secondary incidental findings from genomic sequencing [45], further genetic counselling studies are required in this area to inform academic debate prior to widespread adoption of genomic sequencing.

Adoption by laboratories

The clinical use of whole exome or genome sequencing requires validated, robust, and reproducible methods for sequence generation and analysis. The complexity of high-throughput sequence data does not easily fit within existing laboratory standards for quality management, reference material development, and independent measures of test performance. Additionally, rapid changes and developments in analytical components including chemistry, instrumentation, and bioinformatics protocols present challenges in maintaining best practice whilst simultaneously complying with more rigid regulatory and quality management system requirements and standards [46]. Another challenge for laboratories is the significant increase in the number of identified variants and the scaling of current interpretive practices to meet the increase of variant data. Sharing of experience, data, and interpretation across laboratories is essential for consistent and improved clinical interpretation of variants. Standards for interpretation, access to quality international data, and guidelines for actionable variants will facilitate the uptake of advances in capabilities and quality improvements in pathology services.

In Australia, Europe, and the USA, guidelines are currently under development to equip pathology services with the necessary skills and protocols [44].

Community concerns

Public concern about the misuse of genetic information has been longstanding (http://pathwiki.rcpaqap.com.au/pathwiki/index.php/Introduction), particularly relating to the potential for 'genetic discrimination'. The National Health and Medical Research Council (NHMRC) of Australia defines this as occurring when 'a person, such as an employer, or company treats another person or their relatives differently on the basis of their actual or assumed genetic makeup' (http://www.nhmrc.gov.au/your-health/egenetics/ethics-and-legal-issues/genetic-discrimination). Genetic discrimination is unusual in that it can occur even when a person has not yet developed a condition, if that person's status is assessed as 'at risk'. In Australia, discrimination on the grounds of genetic status is dealt within existing Commonwealth, state, and territory anti-discrimination laws, which generally cover circumstances where discrimination occurs in a public domain such as employment, life insurance, education, or access to other services. In Australia, there is no legislation specifically prohibiting genetic discrimination in particular, unlike, for example, Belgium, France, Germany, and Sweden [47], or the USA.

Commonly, concerns about genetic discrimination focus on access to, and affordability of, health and life insurance. Australia provides an example of a country with established government-funded and private healthcare systems. Although private health insurance premiums in Australia are not determined by personal health status, family history, or genetic test results, life and income insurance applications are assessed on the basis of the applicant's risk. This includes personal medical history, family history, and genetic test results. Applicants for life or income insurance are also expected to disclose participation in studies which return genetic or genomic information to participants. Importantly, therefore, policies governing access of insurance underwriters to results of genomic sequencing can have implications for access to and affordability of life insurance products. However, as it is to be expected that the majority of people have some deleterious variants [13], the governance of their access to results from genomic sequencing in prevention and public health also presents a challenge for insurance underwriters. As a partial response to this challenge, the representative body of Australian life insurance providers (the Investment and Financial Services Association) issued a statement of current industry practice in response to a pilot haemochromatosis screening program stating that, for people who tested positive there will be no impact on their life, disability, and trauma insurance as long as there is no evidence of the disease (http://www.fsc.org.au/downloads/file/MediaReleaseFile/2001_0906_HaemScreenIFSARelease.pdf). This is a useful precedent showing genetic services and the insurance industry working together to develop guidance relating to population screening for a genetic condition.

Genomic information is clearly sensitive, in part because it can be used to predict future risk, although with currently limited accuracy. Another source of sensitivity is the fact that genomic information is shared among family members (see Box 3 for an example of privacy issues in research). A trade-off therefore arises between the need to protect individual privacy and the benefits that can be realised through international data sharing to ensure high quality interpretation about the significance of an individual's genomic profile. Practices and measures that permit confidential data sharing and use whilst protecting confidentiality are needed by pathology services, researchers, policy analysts, and others (see, e.g., [48,49]).

Clearly, there is more research required to understand and quantify the privacy risks in genomic data, and to enable an informed trade-off with access and use for research. Despite this, it is heartening that there have been no major privacy breaches in this area to date.

Box 3. Privacy concerns in genomic research

A recent analysis found that 6.8% of GWAS had the unforeseen potential for misuse by exposing an 'individual research participants' information, including revealing disease status, predicted future likelihood or past presence of other traits, or attempts to link another DNA result with a participant, for example, to determine presence or absence in a research cohort, ancestry, and relatedness (e.g., paternity/non-paternity)' [65]. In another study, Gymrek *et al.* [66] were able to infer the identity of a minority of male participants (12% with false discovery rate of 5%) in public sequencing projects by linking Y chromosome markers with available user-provided information (e.g., surname) from recreational ancestry databases. Additional public databases were used to narrow down the list of possible individuals with surnames segregating with the Y chromosome markers. The information available in public research databases has since been limited to reduce this risk.

Investing in a technological healthcare area

Unlike mortality caused by injuries, genetic disorders, or infectious diseases, premature deaths caused by hazardous lifestyle choices are missed prevention opportunities [50]. Modifiable risk factors such as smoking and poor diet were found to be responsible for 33% of all US deaths in 2000 [51]. Danaei *et al.* [52], who extended the analysis to include other dietary and metabolic risk factors, argue that influencing behaviour to mitigate dietary, lifestyle, and metabolic risk factors can substantially avert more deaths than further improving the healthcare system. Based on the notion that improved lifestyle can reduce morbidity and significantly extend the productive lifespan of individuals, investment in patient education and long-term engagement strategies is likely to have a higher return on investment and save more lives than extending the exclusive investment in traditional health provision (Box 4).

Electronic health and medical records

Currently, most of the health and medical history of patients is collected on paper or in silos of electronic clinical information systems either at the patient's general practitioner or the admitting hospital [53]. Recognising the need for interoperability, many international initiatives are aimed at sharing this information among relevant specialists and to provide patients with the ability to access and integrate their own clinical information [42]. This drives a paradigm shift from evidence-based practice to practice-based evidence where treatment decisions are based on the aggregation of historical patient outcomes with similar clinical presentation, and comparable genetic profiles, rather than a small number of clinical trials where only limited parameters are typically controlled for [54].

Electronic records are a key enabler in the cost-effective development of evidence for practice. For instance, a number of recent initiatives (eMERGE, BioVu) aim to exploit hospital electronic health records (EHRs) for the recruitment of patient cohorts and for undertaking Electronic Health Driven Genome Research (EDGR). The rationale for EDGR is the dramatic efficiencies that can be attained by recruiting patients and mining phenotypic data by searching already existing EHRs. A number of EDGR studies undertaken to date illustrate the spectacular success of this approach [53]. For example, analysis of EHRs coupled with microarray genotyping data from 1317 cases of hypothyroidism and 5053 controls was able to identify four novel genetic variants associated with this disease [55]. Such studies can be undertaken at a fraction (~10%) of the cost of conventionally recruited targeted GWAS.

EDGR also enables phenome-wide association studies (PheWAS). A PheWAS is effectively a reverse GWAS to associate genomic regions with patient phenotypes, a generalisation, which is outside the scope of traditional, dis-

Box 4. Digital products and services

Although much of the information available to individuals will come from medical doctors and health services, social networking shows great potential to enable individuals to self-manage and receive peer support from the community without the intervention of overly stretched health services [67].

Examples include services such as Patients Like Me (PLM), where patients themselves provide information and ideas to one other to manage their diseases [68]. PLM also provides the opportunity for researchers to recruit patients to specific trials, including GWAS. However, while PLM is a specific social network, general services such as Facebook and Twitter can also be utilised to provide health information to consumers. A recent trial used Facebook for a web-based diet survey and recruited a representative cohort for minimal cost [69]. Meanwhile, Twitter feeds can be monitored for signals indicative of disease outbreaks, the progression of the flu season, or reports of adverse drug events [70,71].

There is an increasing trend away from personal computers to mobile devices and the use of applications (Apps) in particular, as the vehicle by which to interact with information from the World Wide Web. Apps provide the opportunity to capture biological data [72] and enable service providers to reach patients, and vice versa, in a ubiquitous manner. Mobile applications for health typically deliver location-based information and interactive tools to record progress and reward health behaviours. Personalised, adaptive technologies that support individual preferences or knowledge and that respond to individual behaviours can be expected to be more powerful than generic one size fits all applications.

An emerging trend in the health technology space is the popularity of personal sensing through mobile phones and standalone commercial products. These generic, cheap, often wearable, devices, that include FitBit, Jawbone, and Nike Fuelband, have the ability to monitor location, movement, heart rate, blood pressure, oxygen levels via skin coloration, temperature, and more throughout the day to provide an opportunity for building personalised models of an individual's health or health needs and detect abnormalities that are predictive of disease. As research has demonstrated that simple physiological measurements, such as heart rate, can be independent predictors of overall health outcome and general mortality [73], there are enormous opportunities to exploit user profiles in targeting information, services, and interactions with individuals to further personalise existing services and increase their impact.

Mobile and telehealth services such as rehabilitation and chronic disease care can also greatly benefit from the use of mobile sensors. Current research is trialling applications with specific patients, such as those with a chronic disease or those undergoing secondary rehabilitation [74,75]. These approaches pave the way for the health service industry to provide whole-of-life care to patients. Increasingly, this approach will move to being preventative in nature [76]. For example, a recent randomised controlled trial showed that secondary cardiac rehabilitation services offered via a mobile phone rather than a clinic showed an 80% increase in uptake and adherence [77].

ease-focused GWAS. One recent example illustrating the utility of this approach mined EHRs from 1290 cases of rheumatoid arthritis with 1236 controls to identify a previously unknown association of autoantibodies with autoimmune risk alleles and clinical diagnoses [56]. Another application of EDGR is to detect drug–gene interactions (pharmacogenomics). For example, to test if concurrent use of tamoxifen and antidepressants (that function by inhibiting cytochrome P450 2D6) is associated with increased breast cancer recurrence, EHRs of 1962 patients using tamoxifen were analysed [57]. The investigators concluded that despite the biological rationale, there was no increased breast cancer recurrence in concomitant use of cytochrome P450 2D6 inhibitors with tamoxifen.

The need for electronic decision tools becomes apparent to support physicians in assessing large volumes of incidental or uncertain genomic information and communicating realistic risk profiles [58]. The potential and benefits of such systems was recently demonstrated in a study on immunosuppressive treatment during organ transplantation [59].

Concluding remarks

Currently, reported cases where an individual's genomic information has played a central role clinically in the diagnosis or treatment of disease is limited to rare genetic diseases. EDGR on larger, better curated study cohorts could enable polygenic predictors of some complex genetic traits on a population level as rapidly as within the next 5 years [10]. With a better understanding of the genomic variability and a library of clinically annotated variants, extending risk prediction to include environmental factors will further improve early detection, diagnosis, and intervention in complex diseases. However, the largest impact on the common multicausal diseases that form the largest burden on the health system will be gained from a whole-of-life approach, where population-based risk stratification forms the basis for ongoing health monitoring and automated personalised health advice. To achieve this, we recommend that research should focus on three priority areas: (i) genotype–phenotype annotation for diagnostic applications; (ii) data interoperability to future-proof health research; and (iii) longitudinal individual health monitoring in clinical trials (Box 5).

Box 5. Outstanding questions

- **What are the relevant genotype–phenotype associations for diagnostic applications?**
  Genomes annotated with clinical information that are shared across hospitals and international research groups will provide the basis for identifying causative variants for rare diseases and improve risk predictors for common disorders. This information helps in developing a personal and precision medicine system that is primarily geared to health optimisation rather than crisis management.
- **What level of data interoperability is needed to future-proof health research?**
  Collecting and providing data from genomic profiling and environmental risk factors needs to be done in a way which is machine readable, uses common language (ontology), and shows relationships between data items. This will ensure future use of this expensive resource in studies with cohort sizes that can cope with the added complexity from multivariate environmental risk factors.
- **Can longitudinal individual health monitoring improve health outcome?**
  New studies are needed to build baseline models that enable the development of personalised early detection warning systems, which will form the foundation of interventions to prevent disease and delivery of whole-of-life healthcare. Careful consideration will need to be given to ensure the effective delivery of personalised risk prediction and stratification of disease screening and prevention based on this risk.

Acknowledgements

The authors would like to thank Naomi R. Wray, Jill Freyne, Christine O'Keefe, Trevor Lockett, and Bernie Tuch for critical reading of the document.

References

1 Bainbridge, M.N. *et al.* (2011) Whole-genome sequencing for optimized patient management. *Sci. Transl. Med.* 3, 87re3
2 Talkowski, M.E. *et al.* (2012) Clinical diagnosis by whole-genome sequencing of a prenatal sample. *N. Engl. J. Med.* 367, 2226–2232
3 Pellatt, A.J. *et al.* (2012) Genetic and lifestyle influence on telomere length and subsequent risk of colon cancer in a case control study. *Int. J. Mol. Epidemiol. Genet.* 3, 184–194
4 Caulfield, T. and McGuire, A.L. (2012) Direct-to-consumer genetic testing: perceptions, problems, and policy responses. *Annu. Rev. Med.* 63, 23–33
5 Bloss, C.S. *et al.* (2013) Impact of direct-to-consumer genomic testing at long term follow-up. *J. Med. Genet.* 50, 393–400
6 Knoppers, B.M. *et al.* (1999) Policy forum: genetic technologies. Commercialization of genetic research and public policy. *Science* 286, 2277–2278
7 Andermann, A. *et al.* (2010) Genetic screening: a conceptual framework for programmes and policy-making. *J. Health Serv. Res. Policy* 15, 90–97
8 Fears, R. and ter Meulen, V. (2013) The perspective from EASAC and FEAM on direct-to-consumer genetic testing for health-related purposes. *Eur. J. Hum. Genet.* 21, 703–707
9 Green, R.C. and Farahany, N.A. (2014) Regulation: the FDA is overcautious on consumer genomics. *Nature* 505, 286–287
10 Chatterjee, N. *et al.* (2013) Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. *Nat. Genet.* 45, 400–405
11 Huber, M. *et al.* (2011) How should we define health? *BMJ* 343, d4163
12 Shashi, V. *et al.* (2014) The utility of the traditional medical genetics diagnostic evaluation in the context of next-generation sequencing for undiagnosed genetic disorders. *Genet. Med.* 16, 176–182
13 Abecasis, G.R. *et al.* (2010) A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073
14 Oetting, W.S. *et al.* (2013) Getting ready for the Human Phenome Project: the 2012 forum of the Human Variome Project. *Hum. Mutat.* 34, 661–666
15 Landrum, M.J. *et al.* (2014) ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* 42, D980–D985
16 Biankin, A.V. and Hudson, T.J. (2011) Somatic variation and cancer: therapies lost in the mix. *Hum. Genet.* 130, 79–91
17 Tamborero, D. *et al.* (2013) Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci. Rep.* 3, 2650
18 Campeau, P.M. *et al.* (2008) Hereditary breast cancer: new genetic developments, new therapeutic avenues. *Hum. Genet.* 124, 31–42
19 So, H-C. *et al.* (2011) Risk prediction of complex diseases from family history and known susceptibility loci, with applications for cancer screening. *Am. J. Hum. Genet.* 88, 548–565
20 Visscher, P.M. *et al.* (2012) Five years of GWAS discovery. *Am. J. Hum. Genet.* 90, 7–24
21 Thomas, D. (2010) Gene–environment-wide association studies: emerging approaches. *Nat. Rev. Genet.* 11, 259–272
22 Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group (2014) The EGAPP initiative: lessons learned. *Genet. Med.* 16, 217–224
23 Patel, C.J. *et al.* (2010) An environment-wide association study (EWAS) on type 2 diabetes mellitus. *PLoS ONE* 5, e10746

24 Ashley, E.A. *et al.* (2010) Clinical assessment incorporating a personal genome. *Lancet* 375, 1525–1535
25 Patel, C.J. *et al.* (2013) Whole genome sequencing in support of wellness and health maintenance. *Genome Med.* 5, 58
26 Rappaport, S.M. and Smith, M.T. (2010) Epidemiology. Environment and disease risks. *Science* 330, 460–461
27 Wild, C.P. (2012) The exposome: from concept to utility. *Int. J. Epidemiol.* 41, 24–32
28 Fenech, M. (2013) A public health genomics approach to improving well-being based on DNA damage prevention using nutritional, lifestyle, environmental and psycho-social strategies. *J. Austral. Coll. Nutr. Environ. Med.* 32, 9–16
29 Polychronakos, C. and Li, Q. (2011) Understanding type 1 diabetes through genetics: advances and prospects. *Nat. Rev. Genet.* 12, 781–792
30 Kussmann, M. and Van Bladeren, P.J. (2011) The extended nutrigenomics – understanding the interplay between the genomes of food, gut microbes, and human host. *Front. Genet.* 2, 21
31 Hivert, M.F. *et al.* (2014) Susceptibility to type 2 diabetes mellitus – from genes to prevention. *Nat. Rev. Endocrinol.* 10, 198–205
32 Minihane, A.M. (2013) The genetic contribution to disease risk and variability in response to diet: where is the hidden heritability? *Proc. Nutr. Soc.* 72, 40–47
33 Fenech, M. *et al.* (2011) Nutrigenetics and nutrigenomics: viewpoints on the current status and applications in nutrition research and practice. *J. Nutrigenet. Nutrigenomics* 4, 69–89
34 Ordovas, J.M. and Corella, D. (2004) Nutritional genomics. *Annu. Rev. Genomics Hum. Genet.* 5, 71–118
35 Kaput, J. *et al.* (2005) The case for strategic international alliances to harness nutritional genomics for public and personal health. *Br. J. Nutr.* 94, 623–632
36 Wilson, C.P. *et al.* (2013) Blood pressure in treated hypertensive individuals with the *MTHFR* 677TT genotype is responsive to intervention with riboflavin: findings of a targeted randomized trial. *Hypertension* 61, 1302–1308
37 Fenech, M.F. (2010) Nutriomes and nutrient arrays – the key to personalised nutrition for DNA damage prevention and cancer growth control. *Genome Integr.* 1, 11
38 van Ommen, B. *et al.* (2010) The Micronutrient Genomics Project: a community-driven knowledge base for micronutrient research. *Genes Nutr.* 5, 285–296
39 Eng, G. *et al.* (2012) Genome technologies and personalized dental medicine. *Oral Dis.* 18, 223–235
40 Caggana, M. *et al.* (2013) Newborn screening: from Guthrie to whole genome sequencing. *Public Health Rep.* 128 (Suppl. 2), 14–19
41 Andermann, A. *et al.* (2011) Guiding policy decisions for genetic screening: developing a systematic and transparent approach. *Public Health Genomics* 14, 9–16
42 Chute, C.G. and Kohane, I.S. (2013) Genomic medicine, health information technology, and patient care. *JAMA* 309, 1467–1468
43 Najafzadeh, M. *et al.* (2012) Barriers to integrating personalized medicine into clinical practice: a best-worst scaling choice experiment. *Genet. Med.* 14, 520–526
44 Rehm, H.L. *et al.* (2013) ACMG clinical laboratory standards for next-generation sequencing. *Genet. Med.* 15, 733–747
45 Shahmirzadi, L. *et al.* (2013) Patient decisions for disclosure of secondary findings among the first 200 individuals undergoing clinical diagnostic exome sequencing. *Genet. Med.* http://dx.doi.org/10.1038/gim.2013.153
46 Gargis, A.S. *et al.* (2012) Assuring the quality of next-generation sequencing in clinical laboratory practice. *Nat. Biotechnol.* 30, 1033–1036
47 Van Hoyweghen, I. and Horstman, K. (2008) European practices of genetic information and insurance: lessons for the Genetic Information Nondiscrimination Act. *JAMA* 300, 326–327
48 Malin, B.A. *et al.* (2013) Biomedical data privacy: problems, perspectives, and recent advances. *J. Am. Med. Inform. Assoc.* 20, 2–6
49 O'Keefe, C.M. and Connolly, C.J. (2010) Privacy and the use of health data for research. *Med. J. Aust.* 193, 537–541
50 McGinnis, J.M. and Foege, W.H. (1993) Actual causes of death in the United States. *JAMA* 270, 2207–2212
51 Mokdad, A.H. *et al.* (2004) Actual causes of death in the United States, 2000. *JAMA* 291, 1238–1245
52 Danaei, G. *et al.* (2009) The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors. *PLoS Med.* 6, e1000058
53 Kohane, I.S. (2011) Using electronic health records to drive discovery in disease genomics. *Nat. Rev. Genet.* 12, 417–428
54 Swisher, A.K. (2010) Practice-based evidence. *Cardiopulm. Phys. Ther. J.* 21, 4
55 Denny, J.C. *et al.* (2011) Variants near FOXE1 are associated with hypothyroidism and other thyroid conditions: using electronic medical records for genome- and phenome-wide studies. *Am. J. Hum. Genet.* 89, 529–542
56 Liao, K.P. *et al.* (2013) Associations of autoantibodies, autoimmune risk alleles, and clinical diagnoses from the electronic medical records in rheumatoid arthritis cases and non-rheumatoid arthritis controls. *Arthritis Rheum.* 65, 571–581
57 Dezentjé, V.O. *et al.* (2010) Effect of concomitant CYP2D6 inhibitor use and tamoxifen adherence on breast cancer recurrence in early-stage breast cancer. *J. Clin. Oncol.* 28, 2423–2429
58 Miller, F.A. *et al.* (2014) Testing personalized medicine: patient and physician expectations of next-generation genomic sequencing in late-stage cancer care. *Eur. J. Hum. Genet.* 22, 391–395
59 Laerum, H. *et al.* (2014) A taste of individualized medicine: physicians' reactions to automated genetic interpretations. *J. Am. Med. Inform. Assoc.* 21, e143–e146
60 Lewis, C.M. and Knight, J. (2012) Introduction to genetic association studies. *Cold Spring Harb. Protoc.* 3, 297–306
61 Eichler, E.E. *et al.* (2010) Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* 11, 446–450
62 Manolio, T.A. *et al.* (2009) Finding the missing heritability of complex diseases. *Nature* 461, 747–753
63 Yang, J. *et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42, 565–569
64 Furrow, R.E. *et al.* (2011) Environment-sensitive epigenetics and the heritability of complex diseases. *Genetics* 189, 1377–1387
65 Johnson, A.D. *et al.* (2011) Temporal trends in results availability from genome-wide association studies. *PLoS Genet.* 7, e1002269
66 Gymrek, M. *et al.* (2013) Identifying personal genomes by surname inference. *Science* 339, 321–324
67 Swan, M. (2012) Crowd-sourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. *J. Med. Internet Res.* 14, e46
68 Frost, J. *et al.* (2011) Patient-reported outcomes as a source of evidence in off-label prescribing: analysis of data from Patients Like Me. *J. Med. Internet Res.* 13, e6
69 Lohse, B. (2013) Facebook is an effective strategy to recruit low-income women to online nutrition education. *J. Nutr. Educ. Behav.* 45, 69–76
70 Lyon, A. *et al.* (2012) Comparison of web-based biosecurity intelligence systems: BioCaster, EpiSPIDER and HealthMap. *Transbound. Emerg. Dis.* 59, 223–232
71 Chew, C. and Eysenbach, G. (2010) Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PLoS ONE* 5, e14118
72 Mosa, A.S.M. *et al.* (2012) A systematic review of healthcare applications for smartphones. *BMC Med. Inform. Decis. Mak.* 12, 67
73 Seccareccia, F. *et al.* (2001) Heart rate as a predictor of mortality: the MATISS project. *Am. J. Public Health* 91, 1258–1263
74 Smith, J.C. and Schatz, B.R. (2010) Feasibility of mobile phone-based management of chronic illness. *AMIA Annu. Symp. Proc.* 2010, 757–761
75 Ryan, D. *et al.* (2012) Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: multicentre randomised controlled trial. *BMJ* 344, e1756
76 Miron-Shatz, T. and Ratzan, S.C. (2011) The potential of an online and mobile health scorecard for preventing chronic disease. *J. Health Commun.* 16 (Suppl. 2), 175–190
77 Varnfield, M. *et al.* (2011) Uptake of a technology-assisted home-care cardiac rehabilitation program. *Med. J. Aust.* 194, S15–S19