Rare disease Roadmap: Navigating the bioinformatics and translational challenges for improved patient health outcomes


PII: S2211-8837(14)00054-9
DOI: http://dx.doi.org/10.1016/j.hlpt.2014.08.007
Reference: HLPT99

To appear in: Health Policy and Technology

Cite this article as: Matthew I. Bellgard, Mark W. Sleeman, Felix Guerrero, Sue Fletcher, Gareth Baynam, Jack Goldblatt, Yaffa Rubinstein, Stephen Groft, Roberto Barrero, Alan H. Bittles, Stephen D. Wilton, Christopher Mason, Tarun Weeramanthri, Rare disease Roadmap: Navigating the bioinformatics and translational challenges for improved patient health outcomes, Health Policy and Technology, http://dx.doi.org/10.1016/j.hlpt.2014.08.007

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Rare Disease Roadmap: Navigating the bioinformatics and translational challenges for improved patient health outcomes

1Matthew I Bellgard, BSc, PhD, 2Mark W Sleeman, PhD, 3Felix Guerrero, PhD, 1,4Sue Fletcher, PhD, 5Gareth Baynam, PhD, 6Jack Goldblatt, PhD, 7Yaffa Rubinstein, PhD, 8Stephen Groft, Pharm D, PhD, 9Roberto Barrero, 1,10Alan H Bittles, MA, PhD, ScD, FRCPath, 1,4Stephen D Wilton PhD, 11Christopher Mason, PhD, 12Tarun Weeramanthri, PhD, FRACP, FAFPHM.

1 Centre for Comparative Genomics, Murdoch University, Western Australia
2 Department of Physiology, Biochemistry & Molecular Biology, Monash University, Victoria, Australia
3 US Department of Agriculture, Agricultural Research Service, Knipling-Bushland US Livestock Insects Research Laboratory, Kerrville, USA.
4 The Australian Neuromuscular Research Institute, Centre for Neuromuscular and Neurological Disorders, The University of Western Australia, Nedlands, Western Australia
5 Genetic Services of Western Australia, King Edward Memorial Hospital. School of Paediatrics and Child Health, University of Western Australia, Institute for Immunology and Infectious Diseases, Murdoch University, Western Australia
6 Director Genetic Services & Familial Cancer Program of Western Australia
7 Office of Rare Disease Research, National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, Maryland, USA
8 National Center for Genome Resources
9 Office of Population Health Genomics, Public Health and Clinical Services Division, Health Department of Western Australia, School of Pathology and Laboratory Medicine, University of Western Australia.
10 Community Genetics, School of Medical Sciences, Edith Cowan University, Western Australia
11 Department of Physiology & Biophysics, Cornell University
12 Clinical Services Division, Health Department of Western Australia

Sources of Support: No funding received

Key words: rare disease, bioinformatics, translational research

Corresponding Author:
Professor M I Bellgard
Email: mbellgard@ccg.murdoch.edu.au
Centre for Comparative Genomics
Murdoch University
South Street
Murdoch WA 6150
Tel: +61 8 9360 6088
Fax: +61 8 9360 7238
Abstract

Rare Disease registries have now been recognized as a global priority for progress in monitoring, documenting the natural course, preventing and treating Rare Diseases. However, a disease registry is only one element of Rare Disease translational research. Here, we outline what we believe are ten key components in comprehensive Rare Disease translational research and describe critical relationships between them. These components are: i) Client-practitioner partnerships; ii) Disease registries; iii) Biobanks; iv) Genomics and other -omics platforms; v) Community-based and population-wide studies; vi) Bioinformatics and high performance computing; vii) Interactions with pharma to facilitate drug discovery; viii) Personalised treatments based on genotype-phenotype correlations; ix) eHealth and a whole of life record; and x) Regulatory frameworks, particularly with regard to specimen and data sharing, and return of results. Each component has its own inherent complexity, but if effectively integrated they will provide for a comprehensive approach to the future management of Rare Diseases, and aid health care providers in delivering services for individuals affected with Rare Diseases. We demonstrate that navigating through the Roadmap can provide relevant health stakeholders with a blueprint to understand the challenges and barriers which need to be overcome within and across the constituent components. The Rare Disease Roadmap will assist decision-making at all health stakeholder levels and enable the seamless integration of new knowledge, standard operating procedures and the implementation of best practice.
Introduction

Our understanding and appreciation of the complexity of the genetic basis of human disease is increasing rapidly. Single gene disorders account for approximately 10,000 human diseases, of which some 7,000 disorders are categorized as Rare Diseases (RDs). There are overlapping operational definitions of RDs and, for example, in Europe, an estimated prevalence of five or fewer affected persons in 10,000 is used. The individual rarity of each individual RD provides inherent challenges for ascertaining epidemiological data and recruiting sufficiently large cohorts for clinical trials or translational research. To address this shortcoming, there is for a need to establish global inclusive RD registries and biobanks (1) (2) (3) (4) that facilitate the coordinated acquisition of fundamental disease information and research specimens to: i) assess the health and economic impact of rare disorders individually or collectively; ii) devise best practice disease management strategies; iii) engage complementary national/international expertise; and iv) build sustainable academic, government and/or industry partnerships; and (v) develop large sustainable resources for translational research. Some challenges and barriers that need to be overcome to facilitate these approaches include the international harmonisation of informed consent, and specimen and data sharing practices.

In 2010, at a workshop on, "Advancing Rare Disease Research: The Intersection of Patient Registries, Biospecimen Repositories, and Clinical Data”, the Office of Rare Disease Research (ORDR) at the National Center for Advancing Translational Sciences (NCATS)/NIH initiated a movement to create a global RD patient registry (2). The workshop produced a set of recommendations, which are being implemented
through the Global Rare Diseases Patient Registry and Data Repository-GRDR-program (see https://grdr.ncats.nih.gov/).

More recently, EURORDIS, NORD and CORD issued a Joint Declaration of 10 Key Principles for Rare Disease Patient Registries (5). This initiative is an important strategic step in enabling the diagnosis and treatment of patients living with a RD. Additional considerations that can complement the principles of the Declaration include how registries interact with orphan drug development (6); personalized and therapeutic interventions (7); RD diagnostics and novel phenotyping strategies (8); population-wide association studies (9); rapidly evolving integrated bioinformatics advances (10); public policy (11); and international standardisation (12). We argue that, to make substantial advances in the translation of RD research to clinical practice all of these activities must coalesce to avoid duplication, and that key conceptual and methodological advances need to be effectively shared and harmonised to promote rapid adaption to the ever-changing RD landscape.

In this paper we identify ten key components necessary for successful RD translational research and development. The ten components are: i) Client-practitioner partnerships; ii) Disease registries; iii) Biobanks; iv) Genomics and other -omics platforms; v) Community-based and population-wide studies; vi) Bioinformatics and high performance computing; vii) Interactions with pharma to facilitate drug discovery; viii) Personalised treatments based on genotype-phenotype correlations; ix) eHealth and a whole of life record; and x) Regulatory frameworks, particularly with regard to specimen and data sharing and return of results. The implementation of each component is extremely complex and all have associated challenges (or noise, to use an engineering analysis) that are technical, logistical, socio-political, ethical, legal, jurisdictional and/or economic in nature, or workforce-related. For instance, a
poorly designed next generation DNA sequencing experiment or incorrect sample extraction from a biobank will introduce technical noise into the research on a given RD (13).

Another example of such a barrier is the limited access to particular RD patient samples from one organisation/jurisdiction that might be overcome via access to large internationally accessible biobanks for the same disease. Thus, the scale of the barrier and its impact can vary within each component and can be mitigated through international harmonization of processes and our improved understanding of biological systems, biotechnological advances, improved partnerships, the introduction of internationally accepted best practice, use of refined bioinformatics analysis methods, effective use of Internet technologies and/or high performance computing. In turn, the adoption of these strategies should ultimately facilitate translational research and ultimately result in improved patient treatment protocols.

The challenge, and the opportunity we see, is to understand how these different components interoperate into a RD Roadmap and to identify and characterize the potential interfaces within and across each component. We discuss how to navigate through this RD Roadmap, identify the challenges faced with each component, as well as discuss the alternative paths that might be taken to ultimately improve the care of individuals living with a RD.

**Ten key components for Translational RD Research**

Ten components regarded as core to translational RD research are outlined in the subsections below and summarised in Figure 1. We describe each of the components under three subheadings; Overview, Challenges and Ways Forward.
1 Patient-Health Professional partnerships

Overview

Individuals living with RDs engage with health professionals to achieve a diagnosis and receive timely and appropriate management. These interactions optimally occur by marrying skill sets to patient need in an environment of trust and open communication. This requires the ability to describe (phenotype) a given scenario in a manner that is captured in a disease- and patient-centric knowledge base. However, this approach is hampered by the lack of scalable and robust RD knowledge management approaches and systems.

Challenges

The major challenge with this component is often the lack of a suitable database/registry that facilitates the capture and accessibility of pertinent patient information. This situation might arise for a number of reasons: a) limited time available to extract or enter relevant patient data; b) delays in obtaining consent to obtain personal and family information; c) limited and restrictive information management systems that either are not user-friendly or are incapable of capturing all of the data elements relevant to effective diagnosis and best practice; d) inability to extract pertinent legacy information from handwritten clinical notes or internal spreadsheets that become a part of defacto laboratory standards; and e) system inability to capture a variety of data from established and evolving diagnostic technologies.

Ways Forward

Mitigation strategies that are currently underway include: a) the establishment of data elements that are common to all RDs (12); b) development of Internet-based
laboratory information management systems that are scalable, secure and operate effectively in a distributed environment (3) (14); c) capture of deep phenotypic information in a structured (ontological) way and strategies to build these structures dynamically (15); d) migration and curation of legacy clinical information systems and filing systems containing important RD patient information into new information management systems; and e) interfaces that mirror clinical workflow to promote timely data capture (16).

If done properly, and at the same time as active research leading to iterative updates of common data elements and ontologies continues to evolve, our knowledge management systems should have the capacity to dynamically and seamlessly incorporate these updates to improve patient care.

2 Disease Registries

Overview

Rare Disease registries represent a core element for effective RD translational research. Both the NIH workshop and the EURORDIS-NORD-CORD Joint Declaration states that patient RD Registries should have a wide geographical if not global scope. The Registries should also focus on a disease or group of diseases, they should show interoperability between RD Patients, use Common Data Elements, and be linked with biobank data (1) (5). Equally importantly, is the requirement for the integration of public-private partnerships within this effort, as the pharmaceutical industry is actively engaged in rare/orphan disease drug development. (6). International patient RD registries are seen to be critical, particularly to the pharmaceutical industry to provide large, easily accessible, cohorts of affected
individuals, and hence there is an impetus for national- and regional-based registries to improve their coordination so that they can feed into international registries.

**Challenges**

Disease registries can suffer from deficient architectural design and inappropriate software technology choices. Integrating legacy systems or new registries must adhere to international standards for common data elements and ontologies, and their scalability. Numerous patient advocate organizations in each country drive research efforts on specific RDs, and the databases often have to vie for limited public and private funds. Sadly, older established registries frequently contain copious amounts of data that have not been curated/validated and so are likely to be discarded or remain unused. The key barrier is how these resources can dynamically capture knowledge about diseases and thus aid in patient management and translational research.

**Ways Forward**

Registries capture data and thus become an information-rich resource for relevant stakeholders. The identification of common data elements and ontologies plays an important role in future-proofing accurate patient disease information for the range of RDs under investigation and a RD Registry Checklist has been proposed (14). There is evidence that pharmaceutical industries are recognizing efficiencies through working together (17) as well as interacting with academia and governments agencies (18). International and national disease registries have become the vehicle with which to initiate private-public dialogues. National peak RD patient advocate bodies are essential to communicate with federal and state-based governments rather than individual disease advocate groups presenting individually. The establishment of
clinical disease registries must have government backing as well as being endorsed and driven by peak RD advocate groups to be sustainable, attract ongoing resourcing and professional and community engagement, and allow for curation.

3 Biobanks

Overview

A plethora of biobank samples are stored in freezers worldwide and while many of these biobanks are recognized and accredited, others are more boutique in nature. Like the registries outlined above, the global push for robust, sustainable biorepositories is urgent and the practices involved have come under increased scrutiny. Internationally agreed standards are needed to ensure high fidelity in biosample extraction, preparation, long-term storage, and recordkeeping. Errors in any of these steps can destroy the utility of valuable samples and data, thus wasting time and resources (13).

Challenges and Barriers

Biobanks can contain samples for genomic, epigenetic, transcriptomic, proteomic, or metabolomic studies. For any biobank, problems can arise due to human errors in sampling, specimen storage, flawed recordkeeping, and inappropriate use of samples. A critical problem is the lack of internationally adopted standard operating procedures (SOPs) and many organisations, both private and public, have their own internally designed SOPs. With increasing knowledge of the effects that sample extraction/preparation/environmental factors have on gene, protein, and metabolite expression, there is a need for well-designed and updated SOPs that reflect best practices in sample handling and storage using current technologies. Questions that arise include: i) reliability of legacy biobanks; ii) type of consent obtained at time of
collection; and iii) procedures for return of results. Other challenges include limiting access to particular biorepositories, limited residual sample amounts, intellectual property, or privacy/competition issues that can inhibit collaborative value-adding studies.

**Ways Forward**

Consistent, internationally acceptable SOPs are important and there is a push to develop this approach world-wide (19). International harmonization of these processes is vital, and technology is greatly improving our understanding of how samples should be managed in biobanks (13) (20).

### 4 Genomics and other -omics platforms

**Overview**

Quantum advances are continually being made in the genomics, transcriptomics, proteomics and metabolomics platform technologies, providing opportunities for significant advances in RD research. For example, whole genome sequencing of patients with a RD, compared to genomes from unaffected individuals, has facilitated the discovery of genetic changes underlying many RDs (21). Integrating the various –omics platforms will make a significant contribution to improved personalized patient care (22).

**Challenges**

Developments in each of the –omics platforms are progressing rapidly and their applicability to RD research is evolving. But as more sequencing platforms have become available to RD researchers, problems with data interpretation have increased. Each sequencing assembly and annotation approach must be optimized, and to blend datasets arising from different approaches requires the development of
new bioinformatics tools, e.g. for the detection of a disease-causing gene mutation (23).

For instance, in metabolomic studies, there are recognized issues associated with specific platforms, e.g. NMR versus mass spectrometry that can have a dramatic impact on sensitivity versus specificity (24). There will also be downstream noise for any comparative analysis between clinical laboratories treating patients with similar phenotypes or therapeutic tracking. Noise can be introduced when incorporating data generated from legacy sequencing technologies in the absence of descriptions of the experimental design. Thus, there are various ways in which noise can enter the system when working with samples targeted for genomic, transcriptomic, proteomic and metabolomic analysis.

**Ways Forward**

Robust Laboratory Information Management Systems (LIMS) are essential for each of the platforms (e.g. an initiative for developing a LIMS for metabolomics studies has commenced in Australia, https://bitbucket.org/ccgmurdoch/mastr-ms/overview). Metadata documentation of experimental design and analysis audit trails must be captured to enable informed decision-making. Bioinformatics tools and systems are central to all of the –omics platforms, not only to provide solutions to data processing, quality control, and analysis but also for data integration.

**5 Population-wide studies**

**Overview**

Genome-wide association studies (GWAS) aim to both identify loci (genetic markers) associated with disease and to locate candidate gene(s) that might become diagnostic and therapeutic targets. With the advances in next generation sequencing, it is now
possible to cost-effectively screen large cohorts with a large number (over a million) of markers (25).

**Challenges**

While GWAS aim to identify common variants, it is recognized that within common diseases, rarer variants may have significant individual effects (9). Within an RD context, the challenge for GWAS is to acquire large cohorts to be able to identify functional rare variants that have a minor allele frequency. There are numerous avenues for noise in population-wide association studies for RDs, and well-documented, inherent problems in GWAS studies include: cohort size, population structure, bioinformatics/biostatistical analysis (26), consideration of environmental factors (27), issues surrounding biobanks (outlined above) and epistasis (gene interaction/phenotype effects) (28).

**Ways Forward**

Large patient cohorts are unlikely in a RD context unless approached on a global scale, including a more detailed understanding of population structure in RD. Many Rare Diseases have a high incidence or are even unique to specific communities and sub-communities, having arisen as founder mutations in past generations. A thorough working knowledge of population genetic structure and sub-structure is therefore paramount, particularly in the many global populations in which endogamous and consanguineous marriage is strongly favoured and Rare Diseases are significantly more likely to be expressed (29). It is unfortunate that, previously, the genomic effects of population stratification have largely been ignored, e.g. with a notable failure to account for caste differences in the design of case-control and cohort studies in India (30). The 1000 Genomes Project will identify variants with a frequency of $\geq$
1% (31) (32) and identify rarer disease susceptible variants, and exome or whole genome studies on patients may help to circumvent the noise inherent in population-wide association studies.

6 Biinformatics and high performance computing

Overview

Biinformatics is involved in devising strategies for the management of data, its analysis and integration with tools that enable rapid scientific discovery, and informed decision-making. Bioinformatics’ issues relate to how patient data are captured and analysed (data standards, experimental design and analysis protocols), diseases are classified (ontologies), data are shared (interoperability, security, privacy), and data are managed (appropriate design and implementation of information management systems). In a RD context, bioinformatics solutions must scale to work across laboratories and jurisdictional boundaries, must be flexible in the transfer of analysis/systems/tools from one RD to another, and modular to integrate new knowledge.

Challenges

There are numerous avenues for noise in bioinformatics. Data currently captured for RDs are managed in many quite disparate formats; laboratory notebooks, spreadsheets and in-house databases containing important phenotypic and epidemiological data, and maintained independently by scientists, clinicians and health workers. They are also often located in geographically dispersed locations, such as offices, institutions, archival repositories and hospitals. This type of challenge is not specific to RD research (33) with information not available in exchangeable data formats, and there is always the potential for groups and individuals to adopt non-scalable data
management practices, to become dependent on restrictive computing resources for storing critical data, and to make poor choices in software development design (14). In addition, there is a world-wide skill shortage for bioinformaticians to ensure that RD research is coordinated globally.

The lack of information systems’ standards to implement a flexible and generic bioinformatics framework that can cope with the ever-evolving advances in technologies, algorithms, data formats and high performance cluster (HPC) infrastructure, while enabling personalized analysis strategies, may pose a challenge to most groups conducting research in the RD domain and other biomedical areas. This bioinformatics framework should be modular, user-friendly and highly customizable to utilize the most comprehensive repertoire of computational workflows and tools to minimize false-negatives and maximize the sensitivity to identify sequence variants associated with human diseases (10, 23).

**Ways Forward**

From a strategic perspective there are a number of international initiatives attempting to address at least some of the bioinformatics bottlenecks in RD research. The European Union has endorsed a Framework 7 project, *RD Connect*, which is an integrated platform connecting databases, registries, biobanks and clinical bioinformatics for RD research (http://rd-connect.eu). There are 27 full partners and 17 associate partners world-wide with guiding principles in security, open source, scalability, ongoing curation, data standards and shared knowledge inter-communicated in an open and collaborative environment. It is possible to make collective decisions on topics such as data management, the types of data that must be captured from each of the RD Roadmap components, data longevity, and where it should be stored and integrated. It is also possible to devise systems that enable non-
computing experts to seamlessly access supercomputing, cloud computing and large-scale storage infrastructure world-wide for both storage and computation while addressing security concerns (10).

The *Yabi* web-based bioinformatics analytical system allows the design and deployment of a flexible and generic bioinformatics framework for specialized data analyses (10). The modular design and architecture of *Yabi* is scalable, highly configurable, user-friendly, secure and open source which facilitates seamless and transparent access to heterogeneous HPC environments. *Yabi* provides an analysis workflow environment that can create and reuse workflows as well as manage large amounts of both raw and processed data in a secure and flexible way across geographically distributed computing resources. These characteristics make *Yabi* an attractive system to coordinate bioinformatics efforts in rare diseases globally.

7 Drug discovery

*Overview*

A host of challenges have always faced the development of safe and effective therapeutics for the treatment of RDs. They include, but are not limited to, difficulties in attracting funding for research, recruiting and design for clinical trials, logistics for manufacture, and regulatory hurdles with national and international agencies. Acknowledging these issues, the US Orphan Drug Act was drafted in 1983 with the aim of increasing the investment of the pharmaceutical/biotechnology industry in the development of new diagnostics, therapeutics, and preventive practice. Prior to its enactment only ten products were approved to treat RD in the United States. However, by 2008 this had increased to over 2000 products obtaining orphan drug designation (34). There have also been significant knock-on effects leading to
increases in the number of clinical trials, which in turn has spurred innovation in novel drug technologies and led to the identification of new disease types. This has, however, not been without issues, such as the prohibitive costs for drugs/therapies for rare or orphan diseases.

Many pharmaceutical and biotechnology companies are now engaged with government, advocacy groups and academic researchers to promote RD research and product development but certain disconnects still exist and are primarily due to a lack of coordinated focus. Despite these shortcomings, optimism is high in this post-genomic era around developing innovative medical approaches, especially for some of the rare conditions that stem from single gene defects or processing errors.

Unfortunately, translational research for RDs, as with any disease, is a resource-intensive undertaking demanding both time and money. All parts of the RD community, whether public or private, must embrace new models of engagement and research directions, such as the use of system-wide datasets, interrogation of biobanks, advanced systems biology and computational technologies from the orphan disorders and clinical trials, with the aim of impacting even the rarest of diseases (35). This can only happen through integrated transparent programmes such as the Rare Disease Roadmap, that share resources and infrastructure to make use of scarce funding, expertise, data, and biological samples.

**Challenges**

Basic research, mostly in academic institutions, has always been instrumental in identifying the molecular mechanisms underlying most RDs and it has accelerated the development of both diagnostics and therapeutics. Obviously the paucity of funding
for such research limits the speed of advances in this field, but additional challenges exist in gaining access to the natural history of RDs, in the development of animal models of RDs, accessing human tissues through biobanks, and access to chemical compound libraries. In the past many of these resource-intensive research and development activities have been conducted by better resourced biotechnology and pharmaceutical companies, even if in isolation from the academic or RD community. The current challenge facing this sector however is the time, cost and risk involved in bringing any new drug to market. It represents a formidable obstacle for pharmaceutical developers and most are cutting their internal R&D efforts. While the issues are multifaceted, they do include increasing number of regulatory hurdles for safety assessments, risk management and post-approval research requirements.

For RDs this is even more devastating, and potential therapies have often languished because of low numbers of patients and difficulties around clinical trial designs and statistical analysis for efficacy studies. As the pharmaceutical industry rationalizes its operations into more integrated network models, we will have to deal with a disconnect that exists between the various stakeholders in RD, namely the pharmaceutical companies, academic institutions, patient groups, and regulatory bodies; these issues include patients’ privacy concerns, intellectual property issues, and patent protection around databases and repositories.

**Ways Forward**

Significant initiatives have been in place to stimulate companies to undertake drug development for RDs, such as research grants, research design support, FDA fee waivers, tax incentives and market exclusivity for orphan drugs, and some promising compounds have been brought to the point of clinical testing. However, many pharmaceutical companies are changing the way they do research and development
and they are rationalizing their operations into more integrated network models that include all stakeholders, including academic research centres, patient groups and public-private partnerships (36). The development of such consortia should allow more funding to drive innovation within particular disease-focus areas such as RDs, and it should leverage the capabilities and expertise of all stakeholders. In such arrangements it is important that registries and biorepositories have common standards, including protection for patients and research participants, and appropriate data-sharing arrangements. The joint development and sharing of chemical compound libraries and new advances in drug-targeting technologies will speed therapeutic development in the RD field. Alignment of objectives to deliver innovative new therapies to treat debilitating disease must be free of intellectual property issues, but retain privacy and avoid conflict of interest of academic groups. Additional elements of an integrated strategy are outlined in the report by the committee on Accelerating Rare Disease Research and Orphan Product Development (37).

8 Personalised treatment

Overview

Personalized treatments are important to RD research in view of their heterogeneous causes. All components in the RD Roadmap facilitate personalised research and therapies are developed that ultimately should lead to better health outcomes. For instance, -omics platforms generate personalized genomic data and GWAS can provide individualised genetic disease risk predictions, that facilitates molecular therapies, such as exon skipping (38), tailored for a specific gene mutation(s). Bioinformatics analysis strategies, tools, integrative analysis, quality of datasets, annotation, curation and evidence-gathering are pivotal (22) to these approaches.
**Challenges**

The major challenge is the difficulty in carrying out trials on small numbers of patients, recruiting and monitoring these patients scattered across large geographic distances and environments, and accounting for individual variation versus treatment effects, thereby strengthening the need for a better knowledge and understanding of population genetic sub-structure. Adverse drug reactions are the fourth largest cause of death in the developed world (39), and drugs are metabolised by a number of pathways, some of which are highly polymorphic systems (40). Different adverse drug reactions can be relatively rare events in the treatment of common diseases, and it will be much harder to link cause and effect when treating RDs. Natural human polymorphisms can influence variation in the natural history of disease progression and pose challenges to track and design appropriate studies.

**Ways Forward**

Well curated genotype and phenotype registries of individuals with RDs will facilitate the investigation of confounding factors in clinical trials, thus allowing tailoring of the drug to the individual and stratifying patient cohorts into likely responders, non-responders and those at risk of a serious adverse drug reaction. New therapies based upon the type of mutation, regardless of the specific gene, are being developed. It has been shown that read-through of nonsense mutations can be achieved with the aminoglycoside antibiotic Gentamicin (41), but this is limited due to the serious side-effects associated with the drug, so extensive research efforts have been undertaken to develop safer read-through compounds, such as Ataluren (42). However, controversy surrounds the mechanism of action and efficacy of this drug (43). Another form of personalised therapy aims to delay the progression of Duchenne muscular dystrophy (DMD) through targeted exon skipping, with antisense oligonucleotides which can
specifically redirect pre-mRNA processing to restore an open reading frame. In this manner, the disease-causing protein truncating mutations are removed and the expression of a shorter but still functional protein is induced. Although not applicable to all genes, antisense oligomer-induced exon skipping may be relevant to genes encoding large structural products such as DMD (44).

The approach of studying the effects of therapies in patient cell models emphasises the importance of quality biobanks of patient cells i.e. cells that are relevant to the disease and appropriate for assessing potential therapies. Improved partnerships between researchers, clinicians and patient groups may lead to an improved shared understanding of the diagnosis, progression and treatment of disease. The International Rare Disease Research Consortium (IRDiRC) has been established to bring together researchers and organisations investing in rare diseases research in order to achieve two main objectives by the year 2020, namely to deliver 200 new therapies for rare diseases and means to diagnose most rare diseases. (http://www.irdirc.org). In future, it may be feasible to examine the natural history of disease in each patient, and look for changes in the ‘slope’ following treatment in N=1 trials (45).

9 eHealth and a whole of life record

Overview

In section 1, the need for enhanced patient/health professional partnerships through registries was discussed. These partnerships extend more broadly to engagement with other stakeholders including the biopharma industry. While it is important to characterize the differences between registries and eHealth personalised records, as part of the RD roadmap we also must understand the conditions under which these
two sources of information can be unified. Ultimately, ongoing patient treatment and the capture of whole-of-life records must be interoperable with leading-edge translational RD research and development. Capturing the natural history of a RD across the patient population is another dimension that must be addressed.

Challenges

There are numerous challenges and barriers to changes in this component, ranging from the simple and accurate data capture of diagnosis and ongoing treatment, and the ability to revise legacy information with more informed disease characterization, to integration of these data sets with national eHealth systems. Training of the next generation of health professionals who specialise in RD should include understanding standardized common data elements and best practice disease ontologies (46). While information management systems can be developed, knowledge gained about a given RD must be transferrable to other rare conditions. Also, Internet-based software systems must be accessible in rural and remote areas where RDs may be expressed in relatively isolated populations, but the Internet bandwidth is reduced.

Ways Forward

Information management systems must be developed to capture knowledge about specific Rare Diseases. This step will enable the generation of generalized rules for application to other rare conditions. Software applications that run on mobile devices that can be used by patients to track their own health need to be developed (47). These solutions must be capable of linkage to eHealth records and utilise agreed standardized vocabularies for common data elements.

10. Policy and Regulatory frameworks

Overview
The key to sustainability is to embed all of the above elements in an integrated and harmonized policy and regulatory framework. Such a framework will ideally be national in scope, international in perspective, and local, regional or community-based in its implementation. While prototype national RD protocols are very useful they only constitute a starting-point to improving the lives of patients with RDs, as each plan will need to consider resourcing, workforce, training, performance indicators and ongoing evaluation. Each country will also have critical differences in its national law with respect to the protection of patient privacy in medical records, disclosure of information such as notifiable disease lists, data collection and storage, research and ethics, orphan drug development, and specimen storage, including biobanks.

**Challenges**

Progressing a RD roadmap requires consideration of jurisdiction-specific regulations and ethical, social and legal structures. Consideration must also be given to education of the public, community advocacy, professional training, the provision of incentives and subsidies, innovative financing (including from the private sector), standard-setting, and streamlining of approval processes where appropriate.

**Ways Forward**

Respectful and meaningful national and international partnerships are the key to introducing and sustaining the change agenda. In democratic countries, consistent advocacy from affected patients and their families expressed through community organizations such as Rare Voices Australia (http://www.rarevoices.org.au) is essential. These groups have a unique capacity to work with motivated individuals in the political arena, health care organizations, government departments, academia and the private sector, to address what are often difficult and complex problems. In the absence of such advocacy, there are many other well-defined and numerically more
pressing health concerns to occupy the attention of policy-makers. With such advocacy and partnerships, data can be assembled from a range of sources, e.g. population level prevalence; incidence and mortality data; quality of life surveys; hospital costing data; personal stories of impact and navigation through existing health care systems; evaluation of new models of care; and clinical trials. Together, they provide a more complete picture of the costs and benefits of moving away from the status quo (11). A comprehensive picture of this nature is required by national and state governments before they will contemplate changes to the regulatory environment.

Discussion

The promise of personalized medicine, through new technologies and new –omics platforms, will only be fully realized through comprehensive system reform and integration, including population-level design elements. The RD Roadmap described in this paper shows that key components of this new paradigm begin and end with the individual living with a RD. There are challenges and opportunities in each of the components, but an even greater challenge in their linkage and integration. The lack of attention and resourcing given to RDs forces a level of analysis and new thinking that is often applicable to more common diseases. The ‘old’ technological and policy challenges relating to disease registries, biobanks, orphan drug development and e-health need addressing as much as the ‘new’ challenges relating to –omics and bioinformatics. Progress in these areas will allow us to realize the potential of the new technologies for both rare and common diseases, but failure to develop a rare disease plan will slow the realization of benefits from personalized medicine.
Acknowledgements

The authors received funding from the Australian National Health and Medical Research Council (APP634485, APP1055319) and EU FP7 Project (HEALTH.2012.2.1.1-1-C): RD Connect: An integrated platform connecting databases, registries, biobanks and clinical bioinformatics for rare disease research. The authors wish to acknowledge their involvement in the International Rare Disease Research Consortium. M Bellgard, initiated this research, via a fellowship under the OECD Co-operative Research Programme: Biological Resource Management for Sustainable Agriculture Systems.
References


Figure 1: Rare Disease Roadmap: key components and relationships between them.
Highlights

We provide a global view of the ten key components necessary in comprehensive Rare Disease translational research and describe critical relationships between them. These components are: i) Client-practitioner partnerships; ii) Disease registries; iii) Biobanks; iv) Genomics and other -omics platforms; v) Community-based and population-wide studies; vi) Bioinformatics and high performance computing; vii) Interactions with pharma to facilitate drug discovery; viii) Personalised treatments based on genotype-phenotype correlations; ix) eHealth and a whole of life record; and x) Regulatory frameworks, particularly with regard to specimen and data sharing, and return of results. We suggest that the Roadmap can provide relevant health stakeholders with a blueprint to understand the challenges and barriers that need to be overcome and that the Rare Disease Roadmap will assist decision-making at all health stakeholder levels and enable the seamless integration of new knowledge, standard operating procedures and the implementation of best practice.