We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,400
Open access books available

117,000
International authors and editors

130M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Biomarkers in Rare Genetic Diseases

Chiara Scotton and Alessandra Ferlini

Abstract

Biomarkers offer a way to speed up medical research by shedding light on the physiopathological mechanisms of disease. Furthermore, biomarkers are considered invaluable tools for monitoring disease progression, prognosis, and response to drugs, especially in clinical trials, where they can be used to assess the efficacy, efficiency, and side effects of novel drugs.

Biomarkers also pave the way to personalised medicine, a rapidly developing field that is of particular interest in rare diseases (RDs), i.e. those with a prevalence of less than 5/10,000, which are often genetic in origin. Although rare genetic diseases may be less appealing targets for pharmaceutical companies, they are nevertheless in urgent need of research into their diagnosis, prevention, treatment, and standards of care.

Here we summarise the state of the art in RDs, genetic diagnosis, and novel strategies aimed at accurately identifying and defining gene mutations, and review the evidence emerging from the latest research and clinical trials. We focus in particular on novel biomarkers, describing the different types discovered so far, highlighting their importance and indicating how they may be translated into research, diagnostics, treatment, and preventative applications in personalised strategies for RDs.

Keywords: biomarker, rare disease, genetic disease, genomics, transcriptomics, proteomics

1. Introduction

As each rare disease (RD) only affects a relatively small number of individuals across the globe, there are often great obstacles to their research, diagnosis, treatment, and prevention. In Europe, a disease is considered to be rare, or orphan, when it affects fewer than 5 people in 10,000, in line with the definitions adopted by the European Committee (EC) in their Orphan
Drugs Regulation N° 141/2000 and Commission Communication COM (2008) 679/2 on RDs: Europe’s challenges [1]. However, RDs are often chronic, progressive, degenerative, life-threatening and/or severely disabling in terms of a patient’s quality of life, frequently leading to a lack or loss of autonomy.

Although infections, allergies, and environmental factors, linked in particular to degenerative and proliferative processes (i.e. auto-immunity or cancer), may be implicated in the onset of RDs, the vast majority, approximately 80%, is caused by genetic defects (though not all RDs are genetic diseases). The signs and symptoms of many RDs may therefore be observed at birth or during childhood, and, indeed, roughly 75% of RDs affect children, including chondrodysplasia, neurofibromatosis, osteogenesis imperfecta, proximal spinal muscular atrophy and Rett syndrome. The RDs that manifest in adulthood, on the contrary, include amyotrophic lateral sclerosis (ALS), and Charcot-Marie-Tooth, Crohn, and Huntington diseases.

Although each individual condition may fit the definition of rare, about 7000 distinct RDs have been identified so far, affecting 6–8% of the global population. In fact, it is estimated that 350 million people worldwide suffer from even rarer conditions, suggesting that one in 20 patients will be affected by an orphan disease [2]. Therefore, collectively, RDs are not at all rare, and as a whole, they generate a considerable socioeconomic burden.

In addition to there being a wide spectrum of RDs, they are also characterised by great variability in the age of onset, signs and symptoms, and patterns of tissue/organ involvement. To further complicate the issue, molecular testing and phenotype analysis reveal that mutations occurring in the same gene can be associated with different clinical diagnoses, and marked intra- and interfamilial phenotype variability has been documented. RDs are therefore often extremely difficult to diagnose, and only about 4000 genes have been identified for the 7000 RDs described in the OMIM database [3]. Understandably, therefore, the IRDiRC [4] has set its members the challenge of diagnosing most, if not all, RDs by 2020, and discovering at least 200 new therapeutic options for their patients.

Nevertheless, without early diagnosis and effective treatment strategies, it is impossible to guarantee any improvement in the quality of life and/or life expectancy of such patients. Furthermore, our lack of knowledge regarding the causes, physiopathological mechanisms, and clinical progression of RDs makes it difficult to apply available treatments and to develop novel therapeutic strategies. In addition, the small number of patients complicates the recruitment of an adequate sample for clinical trials, especially in children, which make up an even smaller percentage of the overall RD population. This is an obvious deterrent to the pharmaceutical industry, which has only limited interest in developing and marketing products for this small consumer base. In order to counter some of these problems, both national and the EU governments have made orphan drug laws and funding a priority, but, despite this recent interest, treatment options are currently only available for 5% of RDs [2].

It is not only RDs that could benefit from more activity in this area, as RD research is also considered pivotal for many common diseases, and has in some cases revealed mechanisms and pathways that have been subsequently associated with other rare or common diseases [5]. Indeed, several RDs have been linked to a high degree of genetic and phenotypic heterogeneity;
for example, mutations occurring in the LMNA gene can cause different disease types by affecting different tissues, such as (i) striated muscle (muscular dystrophy such as Emery-Dreifuss muscular dystrophy and limb-girdle muscular dystrophy or dilated cardiomyopathy), (ii) adipose tissue (lipodystrophy syndromes), (iii) peripheral nerve (peripheral neuropathy such as Charcot-Marie-Tooth disorder) or (iv) accelerated ageing (progeria diseases). There are also clinical signs that can be associated with both genetic and acquired disease. For instance, renal cell carcinoma is characterised by the dysregulation of metabolic pathways (oxygen, iron, and nutrient sensing) which are also manifestations of rare hereditary syndromes such as Von Hippel-Lindau (VHL, OMIM 193300) and Birt-Hogg-Dubé (BHD, OMIM 135150) syndromes, as well as hereditary leiomyomatosis and renal cell carcinoma (HLRCC, OMIM 150800) [5]. It is therefore essential for the research being carried out worldwide to focus on identifying characteristic determinants able to discriminate between specific disease states, stages, and probabilities of responding to particular treatments—put simply, biomarkers.

2. Biomarker: definition and utility

Biomarkers were first described and defined in 2001 by two different review papers [6, 7], both of which suggested that they would be the key to understanding the physiopathology of disease and discovering novel treatment strategies. The classic definition of a biomarker is ‘a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention’. In other words, biomarkers are ‘measurables’ that rely on tools and technologies for assessing body fluids or tissue (blood, urine, cell, skin, etc.), such as DNA analysis [point variants, copy number variation (CNV), translocations, methylation analysis], RNA analysis [expression profile and microRNA (miRNA) characterisation], protein analysis (quantification of circulating proteins), and imaging technologies, or other means of physiological measurement [8].

Figure 1. Literature survey. Number of citations in PubMed in which the keyword 'biomarker' is present in ‘title’ and/or ‘abstract’ from 2001 to 2015.
From their definition, the number of published papers related to biomarker discovery has increased more than 20-fold (Figure 1), and the discovery and development of novel biomarkers have kept pace with technical advances, in particular the advent of high-throughput analysis technologies. Moreover, the large number of grant projects set up over the last 5 years to fund biomarker research, including BIO-NMD [9] and NeurOmics [10], has begun to yield considerable fruits in this field.

The BIO-NMD project is a Europe-wide research network whose aim is to identify and validate biomarkers for rare neuromuscular diseases, such as dystrophinopathies (Becker muscular dystrophy, OMIM 300376; Duchenne muscular dystrophy OMIM 310200; dilated cardiomyopathy, OMIM 302045) and COL6-related myopathies (Bethlem myopathy, OMIM 158810; Ullrich congenital muscular dystrophy, OMIM 254090). Funded by the EU (2009–2012), BIO-NMD set out to investigate different human tissues/cells/fluids using multiple -omic strategies (genomics, transcriptomics, and proteomics), an approach that led to the identification of several biomarkers. Thanks to this project, both plasma and tissue biomarkers that will be useful for monitoring disease progression, prognosis, and treatment response have been described and will ultimately help to pinpoint appropriate options for personalised treatment [11, 12].

In a similar vein, the EU’s NeurOmics project is still ongoing and aims to revolutionise diagnostics and develop new treatments for 10 major neuromuscular and neurodegenerative diseases by using sophisticated -omics technologies. To do this, it has brought together leading European research groups, five highly innovative SMEs, and experts from outside the EU, who are all working to identify genes and develop biomarkers for clinical application, as well as to identify drug targets and improve understanding of the physiopathology of the diseases in question.

This research activity has been largely prompted by the versatility of biomarkers. Indeed, the classical view of biomarkers as a clinical end-point, an objective snapshot that reflects how a patient feels, functions or survives, is extremely reductive. In addition to numerous applications in clinical settings, biomarkers may also serve as a surrogate end-point, a predictor of clinical benefit (or lack thereof) based on epidemiological, therapeutic, physiopathological, or other scientific evidence [13]. In other words, a biomarker may act as a clinically meaningful end-point in clinical trials. Such surrogate end-point biomarkers are foreseeably of particular benefit in RDs, in which a high percentage of diseases without a genetic cause, slow disease progression, chronic nature of the diseases, high heterogeneity of signs and symptoms within the same phenotype, and the difficulty in objectively measuring any change in symptoms dramatically increase the expense of clinical trials. Not only the cost but also the difficulty in undertaking trials based on conventional end-points severely curtails their number, and the lack of sensitive, specific, and timely outcome measures hinders the discovery and development of novel treatments.

However, a biomarker can lessen the burden of the clinical trial process by providing information about the safety and efficacy of treatments before the collection of definitive clinical data, which provides the opportunity for mid-course re-appraisal, and even interruption if the intervention being investigated is revealed as potentially harmful to participants [14]. Indeed,
Biomarkers are far superior to subjective measurements, which may not be directly associated to a disease characteristic, or able to detect small changes, especially in the short term. Biomarkers, on the other hand, can provide an objective measurement of aspects precisely correlated to a specific disease condition, potentially enabling small changes in status to be identified, the disease progression to be assessed, and the likely effects of the therapeutic intervention to be predicted [1] while the trial is ongoing, as well as in real-world settings.

Considering the versatility of biomarkers, the European Medicines Agency (EMA) has attempted to standardise them by drawing up a list of the features of an ‘ideal’ biomarker, namely [15]:

- **Analytical validity.** Like a fingerprint, a biomarker should enable measurement, within a specific range, of a parameter able to accurately and clearly distinguish between altered/normal status or treatment response/non-response. The test(s) used to detect a biomarker should be accurate, reliable, and reproducible, and their technological limits clearly defined. As the analytical accuracy depends on laboratory procedures, such as sample preparation and technology application, these should be reported in order to ensure reproducibility of biomarker discovery and validation.

- **Clinical validity.** Like a mirror, a biomarker should accurately reflect the features of a disease (or treatment), detecting even small changes, and not be influenced by circumstantial factors such as diet, exercise, stress, age, sex, or the environment, i.e. an alteration in the disease features should always be reflected by the biomarker, and a difference in the biomarker should always reflect a change in the disease. In other words, an ideal biomarker will identify specific disease parameters and be sensitive to any change in them. Likewise, to be clinically valid, a biomarker must display a high degree of accuracy (indicating correctly whether a patient has or does not have the disease or treatment effects in the vast majority of cases).

- **Clinical utility.** Like a prophet, a biomarker should herald the outcome of a given situation/intervention. In other words, biomarkers should predict the members of a population who will develop a disease, manifest a disease progression, or respond to a specific treatment. The clinical utility of a biomarker in an appropriate population can be measured by two predictive values, the PPV and NPV, which are respectively used to quantify the probability that a person with a positive test for that biomarker will manifest the outcome predicted by the test, and the probability that a person with negative test will not respond to the intervention/treatment.

- **Non-invasiveness.** Like an open door, a biomarker should grant accessibility, i.e. enable an early, sensitive measurement, of disease severity etc., via the simple collection of body fluids (urine/blood) or scanned images (e.g. MRI or PET, etc.). This will allow a disorder to be monitored at different time points, without recourse to invasive procedures such as biopsies or tissue analysis.

- **Feasibility.** Like the passage of time, a biomarker should be practical to identify and measure, as well as invariable, irrespective of the type of sample collection, processing procedures, or methods used in its detection.
• Time and cost-effectiveness. Like a moneybox, a biomarker should be quick and easy to use and not be so expensive and time-consuming to measure that it cannot be used as a surrogate endpoint in clinical trials or to aid diagnostics and disease monitoring.

Biomarkers that possess all these features will inevitably lead to improvements in clinical trials, especially in the field of personalised medicine. Personalised medicine shifts the current ‘one-size-fits-all’ approach to a more individual line of attack or defence, centred on giving ‘the right drug to the right patients at the right time’ [16]. This is particularly crucial in RDs, in which successful treatment development is generally hindered by the small number of patients and short runs that characterise trials for novel interventions.

3. Strategies for biomarker discovery

In recent years, novel techniques and strategies have emerged for biomarker discovery, and there are currently two major approaches being applied:

• **Candidate approach.** This is a hypothesis-driven method based on knowledge of the relevant physiopathological processes, disease pathway(s), or key molecule(s). It analyses the known gene/protein and their linking products in order to discover a qualitative or quantitative variation in diseased samples (fluids, cells, tissues) with respect to normal ones.

• **High-throughput approach.** This is a hypothesis-free strategy that takes advantage of the development of novel techniques for generating very large amounts of data to compare pathological and normal status. This ‘big-data’ approach is extremely powerful, although it is cost-intensive and requires significant time for validation and clinical definition of the biomarkers identified.

3.1. Discovery of genetic variations

Next-generation sequencing (NGS) techniques are based on high-throughput genomic and transcriptomic sequencing. In brief, target regions can be isolated from the entire genome by hybridisation to complementary sequences. This ‘capturing’ is performed on demand, to isolate sequences that may consist of protein-coding regions only (whole exome sequencing), a specifically targeted gene region (focusing on a limited number of known genes), or the entire genome (whole-genome sequencing). The captured region can then be sequenced by one of several methods (pyrosequencing, 454 Roche; sequencing by reversible termination, Illumina; sequencing by ligation, Solid; semiconductor sequencing, Ion Torrent), and the resulting output is composed of several sequence reads, which are then computationally aligned to the known genome in order to unravel any variations, such as small insertions or deletions [17]. Unlike traditional Sanger sequencing, which reads a sequence base by base, NGS is very time-efficient, enabling the simultaneous analysis of millions of base pairs organised in multiple aligned reads. Despite its efficiency, however, NGS is unable to detect dynamic mutations (e.g. triplet expansions) and still has limited capability to identify CNVs. Nevertheless, while we await the development of specific algorithms to overcome these limitations, NGS can be
As an example, Martell et al. have provided a clear indication of biomarker accessibility and utility in Morquio A syndrome, also named mucopolysaccharidosis IVA (MPS, OMIM: 253000, prevalence: 1–9/1,000,000). This recessive lysosomal storage disorder is caused by a mutation in N‐acetylgalactosamine‐6‐sulfatase gene (GALNS), which codes for keratan sulphate and chondroitin‐6‐sulphate. The mutation results in a wide spectrum of clinical features involving skeletal, cardiac, pulmonary, corneal, and hearing impairment, and the identification of biomarkers able to monitor the response to enzyme replacement therapy during clinical trials is long past due. To this end, the authors measured the plasma levels of 88 candidate proteins, finding that three of them (alpha‐1‐antitrypsin, lipoprotein a, and serum amyloid P) may be suitable surrogate end‐points for clinical trials [31].

The main advantage of techniques that can assess biomarkers in body fluids is, of course, their lack of invasiveness. In this regard, a new protein technology, the SOMA scan assay—an aptamer‐based method able to recognise specific protein epitopes—has been used to evaluate protein levels in the sera of DMD patients. By using this technology to compare serum samples from two independent DMD cohorts with healthy individuals, 44 serum biomarkers were identified [32]. Similarly, Auray‐Blais et al. have recently applied novel MS‐based high‐throughput technologies to protein biomarker discovery in the urine samples of patients affected by Fabry disease, succeeding in identifying the lyso‐Gb3/related analogue profile as a diagnostic biomarker [33].

Low invasiveness is also a feature of the most commonly used method of measuring and validating protein biomarkers, the immunoassay. Immunoassays are based on the ability of monoclonal antibodies to capture and detect specific protein domains and enable the simultaneous investigation of several proteins using very low amounts of samples. For example, in idiopathic pulmonary fibrosis (IPF, OMIM 178500), a rare lethal lung disease (prevalence: 1–5/10,000) of unknown aetiology and variable and unpredictable course, a multiplexed assay has been used to simultaneously evaluate 92 proteins in plasma samples from more than 200 patients. By these means, three biomarkers predictive of IPF outcome were identified [34].

Other studies have used the ELISA immunoassay to evaluate serum levels of an extracellular matrix glycoprotein, tenascin‐C (TN‐C), in Emery‐Dreifuss muscular dystrophy (EMD, OMIM 310300), a rare neuromuscular disorder (1–9/1,000,000) characterised by muscular weakness and atrophy, with early joint contractures and cardiomyopathy, finding an association between elevated circulating TN‐C levels and an increased risk of developing dilated cardiomyopathy [35].

Due to the low invasiveness of the methods involved, proteomic biomarkers are also very appealing as surrogate end‐points in clinical trials and/or screening (e.g. neonatal testing).
As an example, Martell et al. have provided a clear indication of biomarker accessibility and utility in Morquio A syndrome, also named mucopolysaccharidosis IVA (MPS, OMIM: 253000, prevalence: 1–9/1,000,000). This recessive lysosomal storage disorder is caused by a mutation in N‐acetylgalactosamine‐6‐sulfatase gene (GALNS), which codes for keratan sulphate and chondroitin‐6‐sulphate. The mutation results in a wide spectrum of clinical features involving skeletal, cardiac, pulmonary, corneal, and hearing impairment, and the identification of biomarkers able to monitor the response to enzyme replacement therapy during clinical trials is long past due. To this end, the authors measured the plasma levels of 88 candidate proteins, finding that three of them (alpha‐1‐antitrypsin, lipoprotein a, and serum amyloid P) may be suitable surrogate end‐points for clinical trials [31].

The main advantage of techniques that can assess biomarkers in body fluids is, of course, their lack of invasiveness. In this regard, a new protein technology, the SOMA scan assay—an aptamer‐based method able to recognise specific protein epitopes—has been used to evaluate protein levels in the sera of DMD patients. By using this technology to compare serum samples from two independent DMD cohorts with healthy individuals, 44 serum biomarkers were identified [32]. Similarly, Auray‐Blais et al. have recently applied novel MS‐based high‐throughput technologies to protein biomarker discovery in the urine samples of patients affected by Fabry disease, succeeding in identifying the lyso‐Gb3/related analogue profile as a diagnostic biomarker [33].

Low invasiveness is also a feature of the most commonly used method of measuring and validating protein biomarkers, the immunoassay. Immunoassays are based on the ability of monoclonal antibodies to capture and detect specific protein domains and enable the simultaneous investigation of several proteins using very low amounts of samples. For example, in idiopathic pulmonary fibrosis (IPF, OMIM 178500), a rare lethal lung disease (prevalence: 1–5/10,000) of unknown aetiology and variable and unpredictable course, a multiplexed assay has been used to simultaneously evaluate 92 proteins in plasma samples from more than 200 patients. By these means, three biomarkers predictive of IPF outcome were identified [34]. Other studies have used the ELISA immunoassay to evaluate serum levels of an extracellular matrix glycoprotein, tenascin‐C (TN‐C), in Emery‐Dreifuss muscular dystrophy (EMD, OMIM 310300), a rare neuromuscular disorder (1–9/1,000,000) characterised by muscular weakness and atrophy, with early joint contractures and cardiomyopathy, finding an association between elevated circulating TN‐C levels and an increased risk of developing dilated cardiomyopathy [35].

Due to the low invasiveness of the methods involved, proteomic biomarkers are also very appealing as surrogate end‐points in clinical trials and/or screening (e.g. neonatal testing).

4.3. Other biomarkers

As mentioned earlier, imaging technologies, and indeed any diagnostic test that is able to measure the disease status in patients, are useful for measuring, and therefore for investigating certain biomarkers. Magnetic resonance imaging (MRI), for example, is a safe and non‐invasive method of analysing muscle, connective tissue, fat, and bone. Indeed, Kinali and co‐workers have demonstrated that the MRI scan, focused on particular muscles, can serve as a biomarker.

Role of Biomarkers in Medicine 240
As an example, Martell et al. have provided a clear indication of biomarker accessibility and utility in Morquio A syndrome, also named mucopolysaccharidosis IVA (MPS, OMIM: 253000, prevalence: 1–9/1,000,000). This recessive lysosomal storage disorder is caused by a mutation in N‐acetylgalactosamine‐6‐sulfatase gene (GALNS), which codes for keratan sulphate and chondroitin‐6‐sulphate. The mutation results in a wide spectrum of clinical features involving skeletal, cardiac, pulmonary, corneal, and hearing impairment, and the identification of biomarkers able to monitor the response to enzyme replacement therapy during clinical trials is long past due. To this end, the authors measured the plasma levels of 88 candidate proteins, finding that three of them (alpha‐1‐antitrypsin, lipoprotein a, and serum amyloid P) may be suitable surrogate end‐points for clinical trials [31].

The main advantage of techniques that can assess biomarkers in body fluids is, of course, their lack of invasiveness. In this regard, a new protein technology, the SOMA scan assay—an aptamer‐based method able to recognise specific protein epitopes—has been used to evaluate protein levels in the sera of DMD patients. By using this technology to compare serum samples from two independent DMD cohorts with healthy individuals, 44 serum biomarkers were identified [32]. Similarly, Auray‐Blais et al. have recently applied novel MS‐based high‐throughput technologies to protein biomarker discovery in the urine samples of patients affected by Fabry disease, succeeding in identifying the lyso‐Gb3/related analogue profile as a diagnostic biomarker [33].

Low invasiveness is also a feature of the most commonly used method of measuring and validating protein biomarkers, the immunoassay. Immunoassays are based on the ability of monoclonal antibodies to capture and detect specific protein domains and enable the simultaneous investigation of several proteins using very low amounts of samples. For example, in idiopathic pulmonary fibrosis (IPF, OMIM 178500), a rare lethal lung disease (prevalence: 1–5/10,000) of unknown aetiology and variable and unpredictable course, a multiplexed assay has been used to simultaneously evaluate 92 proteins in plasma samples from more than 200 patients. By these means, three biomarkers predictive of IPF outcome were identified [34]. Other studies have used the ELISA immunoassay to evaluate serum levels of an extracellular matrix glycoprotein, tenascin‐C (TN‐C), in Emery‐Dreifuss muscular dystrophy (EMD, OMIM 310300), a rare neuromuscular disorder (1–9/1,000,000) characterised by muscular weakness and atrophy, with early joint contractures and cardiomyopathy, finding an association between elevated circulating TN‐C levels and an increased risk of developing dilated cardiomyopathy [35].

Due to the low invasiveness of the methods involved, proteomic biomarkers are also very appealing as surrogate end‐points in clinical trials and/or screening (e.g. neonatal testing).

4.3. Other biomarkers

As mentioned earlier, imaging technologies, and indeed any diagnostic test that is able to measure the disease status in patients, are useful for measuring, and therefore for investigating certain biomarkers. Magnetic resonance imaging (MRI), for example, is a safe and non‐invasive method of analysing muscle, connective tissue, fat, and bone. Indeed, Kinali and co‐workers have demonstrated that the MRI scan, focused on particular muscles, can serve as a biomarker.
As an example, Martell et al. have provided a clear indication of biomarker accessibility and utility in Morquio A syndrome, also named mucopolysaccharidosis IVA (MPS, OMIM: 253000, prevalence: 1–9/1,000,000). This recessive lysosomal storage disorder is caused by a mutation in N‐acetylgalactosamine‐6‐sulfatase gene (GALNS), which codes for keratan sulphate and chondroitin‐6‐sulphate. The mutation results in a wide spectrum of clinical features involving skeletal, cardiac, pulmonary, corneal, and hearing impairment, and the identification of biomarkers able to monitor the response to enzyme replacement therapy during clinical trials is long past due. To this end, the authors measured the plasma levels of 88 candidate proteins, finding that three of them (alpha‐1‐antitrypsin, lipoprotein a, and serum amyloid P) may be suitable surrogate end‐points for clinical trials [31].

The main advantage of techniques that can assess biomarkers in body fluids is, of course, their lack of invasiveness. In this regard, a new protein technology, the SOMA scan assay—an aptamer‐based method able to recognise specific protein epitopes—has been used to evaluate protein levels in the sera of DMD patients. By using this technology to compare serum samples from two independent DMD cohorts with healthy individuals, 44 serum biomarkers were identified [32]. Similarly, Auray‐Blais et al. have recently applied novel MS‐based high‐throughput technologies to protein biomarker discovery in the urine samples of patients affected by Fabry disease, succeeding in identifying the lyso‐Gb3/related analogue profile as a diagnostic biomarker [33].

Low invasiveness is also a feature of the most commonly used method of measuring and validating protein biomarkers, the immunoassay. Immunoassays are based on the ability of monoclonal antibodies to capture and detect specific protein domains and enable the simultaneous investigation of several proteins using very low amounts of samples. For example, in idiopathic pulmonary fibrosis (IPF, OMIM 178500), a rare lethal lung disease (prevalence: 1–5/10,000) of unknown aetiology and variable and unpredictable course, a multiplexed assay has been used to simultaneously evaluate 92 proteins in plasma samples from more than 200 patients. By these means, three biomarkers predictive of IPF outcome were identified [34]. Other studies have used the ELISA immunoassay to evaluate serum levels of an extracellular matrix glycoprotein, tenascin‐C (TN‐C), in Emery‐Dreifuss muscular dystrophy (EMD, OMIM 310300), a rare neuromuscular disorder (1–9/1,000,000) characterised by muscular weakness and atrophy, with early joint contractures and cardiomyopathy, finding an association between elevated circulating TN‐C levels and an increased risk of developing dilated cardiomyopathy [35].

Due to the low invasiveness of the methods involved, proteomic biomarkers are also very appealing as surrogate end‐points in clinical trials and/or screening (e.g. neonatal testing).

4.3. Other biomarkers

As mentioned earlier, imaging technologies, and indeed any diagnostic test that is able to measure the disease status in patients, are useful for measuring, and therefore for investigating certain biomarkers. Magnetic resonance imaging (MRI), for example, is a safe and non‐invasive method of analysing muscle, connective tissue, fat, and bone. Indeed, Kinali and co‐workers have demonstrated that the MRI scan, focused on particular muscles, can serve as a biomarker.
As an example, Martell et al. have provided a clear indication of biomarker accessibility and utility in Morquio A syndrome, also named mucopolysaccharidosis IVA (MPS, OMIM: 253000, prevalence: 1–9/1,000,000). This recessive lysosomal storage disorder is caused by a mutation in N‐acetylgalactosamine‐6‐sulfatase gene (GALNS), which codes for keratan sulphate and chondroitin‐6‐sulphate. The mutation results in a wide spectrum of clinical features involving skeletal, cardiac, pulmonary, corneal, and hearing impairment, and the identification of biomarkers able to monitor the response to enzyme replacement therapy during clinical trials is long past due. To this end, the authors measured the plasma levels of 88 candidate proteins, finding that three of them (alpha‐1‐antitrypsin, lipoprotein a, and serum amyloid P) may be suitable surrogate end‐points for clinical trials [31].

The main advantage of techniques that can assess biomarkers in body fluids is, of course, their lack of invasiveness. In this regard, a new protein technology, the SOMA scan assay—an aptamer‐based method able to recognise specific protein epitopes—has been used to evaluate protein levels in the sera of DMD patients. By using this technology to compare serum samples from two independent DMD cohorts with healthy individuals, 44 serum biomarkers were identified [32]. Similarly, Auray‐Blais et al. have recently applied novel MS‐based high‐throughput technologies to protein biomarker discovery in the urine samples of patients affected by Fabry disease, succeeding in identifying the lyso‐Gb3/related analogue profile as a diagnostic biomarker [33].

Low invasiveness is also a feature of the most commonly used method of measuring and validating protein biomarkers, the immunoassay. Immunoassays are based on the ability of monoclonal antibodies to capture and detect specific protein domains and enable the simultaneous investigation of several proteins using very low amounts of samples. For example, in idiopathic pulmonary fibrosis (IPF, OMIM 178500), a rare lethal lung disease (prevalence: 1–5/10,000) of unknown aetiology and variable and unpredictable course, a multiplexed assay has been used to simultaneously evaluate 92 proteins in plasma samples from more than 200 patients. By these means, three biomarkers predictive of IPF outcome were identified [34]. Other studies have used the ELISA immunoassay to evaluate serum levels of an extracellular matrix glycoprotein, tenascin‐C (TN‐C), in Emery‐Dreifuss muscular dystrophy (EMD, OMIM 310300), a rare neuromuscular disorder (1–9/1,000,000) characterised by muscular weakness and atrophy, with early joint contractures and cardiomyopathy, finding an association between elevated circulating TN‐C levels and an increased risk of developing dilated cardiomyopathy [35].

Due to the low invasiveness of the methods involved, proteomic biomarkers are also very appealing as surrogate end‐points in clinical trials and/or screening (e.g. neonatal testing).

4.3. Other biomarkers

As mentioned earlier, imaging technologies, and indeed any diagnostic test that is able to measure the disease status in patients, are useful for measuring, and therefore for investigating certain biomarkers. Magnetic resonance imaging (MRI), for example, is a safe and non‐invasive method of analysing muscle, connective tissue, fat, and bone. Indeed, Kinali and co‐workers have demonstrated that the MRI scan, focused on particular muscles, can serve as a biomarker.
As an example, Martell et al. have provided a clear indication of biomarker accessibility and utility in Morquio A syndrome, also named mucopolysaccharidosis IVA (MPS, OMIM: 253000, prevalence: 1–9/1,000,000). This recessive lysosomal storage disorder is caused by a mutation in N-acetylgalactosamine-6-sulfatase gene (GALNS), which codes for keratan sulphate and chondroitin-6-sulphate. The mutation results in a wide spectrum of clinical features involving skeletal, cardiac, pulmonary, corneal, and hearing impairment, and the identification of biomarkers able to monitor the response to enzyme replacement therapy during clinical trials is long past due. To this end, the authors measured the plasma levels of 88 candidate proteins, finding that three of them (alpha-1-antitrypsin, lipoprotein a, and serum amyloid P) may be suitable surrogate end-points for clinical trials [31].

The main advantage of techniques that can assess biomarkers in body fluids is, of course, their lack of invasiveness. In this regard, a new protein technology, the SOMA scan assay—an aptamer-based method able to recognise specific protein epitopes—has been used to evaluate protein levels in the sera of DMD patients. By using this technology to compare serum samples from two independent DMD cohorts with healthy individuals, 44 serum biomarkers were identified [32]. Similarly, Auray-Blais et al. have recently applied novel MS-based high-throughput technologies to protein biomarker discovery in the urine samples of patients affected by Fabry disease, succeeding in identifying the lyso-Gb3/related analogue profile as a diagnostic biomarker [33].

Low invasiveness is also a feature of the most commonly used method of measuring and validating protein biomarkers, the immunoassay. Immunoassays are based on the ability of monoclonal antibodies to capture and detect specific protein domains and enable the simultaneous investigation of several proteins using very low amounts of samples. For example, in idiopathic pulmonary fibrosis (IPF, OMIM 178500), a rare lethal lung disease (prevalence: 1–5/10,000) of unknown aetiology and variable and unpredictable course, a multiplexed assay has been used to simultaneously evaluate 92 proteins in plasma samples from more than 200 patients. By these means, three biomarkers predictive of IPF outcome were identified [34]. Other studies have used the ELISA immunoassay to evaluate serum levels of an extracellular matrix glycoprotein, tenascin-C (TN-C), in Emery-Dreifuss muscular dystrophy (EMD, OMIM 310300), a rare neuromuscular disorder (1–9/1,000,000) characterised by muscular weakness and atrophy, with early joint contractures and cardiomyopathy, finding an association between elevated circulating TN-C levels and an increased risk of developing dilated cardiomyopathy [35].

Due to the low invasiveness of the methods involved, proteomic biomarkers are also very appealing as surrogate end-points in clinical trials and/or screening (e.g. neonatal testing).

4.3. Other biomarkers

As mentioned earlier, imaging technologies, and indeed any diagnostic test that is able to measure the disease status in patients, are useful for measuring, and therefore for investigating certain biomarkers. Magnetic resonance imaging (MRI), for example, is a safe and non-invasive method of analysing muscle, connective tissue, fat, and bone. Indeed, Kinali and co-workers have demonstrated that the MRI scan, focused on particular muscles, can serve as a biomarker...
for disease progression in Duchenne muscular dystrophy (DMD, OMIM: 310200), a rare neuromuscular disease (affecting 1/3300 male births) characterised by rapidly progressive muscle weakness and wasting due to degeneration of skeletal, smooth, and cardiac muscles. MRI can be used to accurately identify which type of muscles is sufficiently preserved in DMD, making it a reliable tool for use in clinical trials. Similarly, MRI scans of muscle biopsies are currently being used to correlate the clinical features of muscle diseases with the structure and morphology of muscle fibres [36].

Neurophysiological measurements can also be exploited as imaging biomarkers. For instance, Vucic and colleagues have reported that transcranial magnetic stimulation (TMS) is a useful and non-invasive method of assessing the functional integrity of the motor cortex and its corticomotoneuronal projections in ALS. Despite their similarities, TMS was able to reliably distinguish between ALS and similar peripheral disorders, thereby demonstrating its potential diagnostic utility [37].

In fact, imaging biomarkers are generally considered very appealing, generating a large amount of intensive research in recent years. The ultimate aim of such research is the development of innovative methods of using imaging tools for the detection and monitoring of the signs and symptoms of RDs.

5. Applications and clinical translation of biomarkers

5.1. Diagnostic/prognostic biomarkers

A diagnostic, or prognostic, biomarker is one that identifies a disease or quantifies its pathogenic factors (Figure 4). Essentially, they are signatures that divide the population into healthy and diseased individuals, but in some cases they can finely stratify the disease phenotype into different degrees of severity or sub-phenotypes. The routine diagnostic markers classically used in clinical practice are temperature, blood pressure, and cholesterol levels, among others, whereas in genetic diseases, according to the IRDiRC statement [4], all gene mutations known to cause a Mendelian disease have to be considered their primary genetic biomarkers. For example, DMD, the most common fatal genetic disorder diagnosed during early childhood, arises through mutations in the causative dystrophin (DMD) gene, which are therefore considered disease biomarkers, and can accordingly be used to select patients for enrolment in clinical trials [38].

In some cases, mutations in causative genes can be considered biomarkers of disease severity. This is the case in fragile X syndrome (FXS, OMIM: 300624), a rare intellectual disability disorder with an estimated prevalence about 1 in 2500 to 5000 men and 1 in 4000 to 6000 women. FXS is caused by an expanded CGG triple-repeat located within the 5' UTR of the FMR1 gene. The triplet expansion variability defines four different phenotypes, ranging from healthy to a severe phenotype, and can therefore be used to distinguish between them [39].

In ALS (OMIM 105400), the situation is less clear cut. ALS is a devastating neurodegenerative disease with an incidence of 1/50,000 per year. Although several mutated genes have been
identified in ALS (DCTN1, OMIM 601143; PRPH, OMIM 170710; SOD1 OMIM 147450; NEFH OMIM 162230), the vast majority of patients do not show a defined genetic defect. This would seem to indicate that the causative gene is still missing [40], and research in this area has therefore focused on the discovery of specific biomarkers able to assist clinical diagnosis and monitor the disease progression. In this regard, Hwang et al. have correlated an increased level of HMGB1, non-histone architectural protein, in serum samples with the onset of ALS, even in early stages of the disease. This increased level of HMGB1 could also be useful as a severity biomarker, since they also found higher HMGB1 levels in patients with a severe disease status [41]. Moreover, the same group has recently correlated a reduction in the protein level of LG72 gene, activator of D-amino acid oxidase, to the pathogenesis of ALS [42].

Another example of a diagnostic biomarker has been found in Alexander disease (ALXDRD, OMIM: 203450), a very rare neurodegenerative disorder (incidence of 1/2.7 million per year) characterised by varying degrees of macrocephaly, spasticity, ataxia, and seizures. It ultimately leads to psychomotor regression and death, and causative mutations have been identified in Glial Fibrillary Acidic Protein (GFAP), the major intermediate filament protein of astrocytes, which result in toxic accumulation of the protein. Animal model studies have demonstrated that transactivation of the GFAP promoter is an early indicator of the disease process, and that GFAP level in the CSF could be a potential biomarker in human patients [43].

Biomarkers used in clinical practice to improve disease progression monitoring or disease-risk prediction are defined as prognostic. Simply put, a prognostic biomarker provides information on the course of a disease in an untreated individual, and an example has been identified for Marfan syndrome (MFS, OMIM: 154700), a systemic disease of the connective tissue characterised by a wide spectrum of cardiovascular, skeletal muscular, ophthalmic, and pulmonary manifestations. With an estimated prevalence of around 1/5000, patients affected by MFS suffer from an increased risk of cardiovascular complications that lead to premature death, and a correlation has been demonstrated between the larger aortic root diameters, coupled to a faster aortic root growth, and high serum levels of transforming growth factor-β (TGF-β). Increasing levels of TGF-β predict cardiovascular events and thereby possesses significant prognostic value [44].

Another biomarker for cardiac muscular involvement has been found in Fabry disease (FD, OMIM: 301500), a rare systematic disease (prevalence 1–5/10 000) characterised by the
accumulation of globotriaosylceramide in the plasma and cellular lysosomes of vessels, nerves, tissues, and organs throughout the body. This accumulation leads to progressive skin lesions, renal failure, cardiac and cerebrovascular involvement, and peripheral neuropathy. Continuously elevated cardiac troponin I (cTNI), a laboratory parameter well known to reflect acute and chronic cardiac muscle damage, has been demonstrated in a substantial proportion of patients with FD, suggesting that raised cTNI levels could be a useful laboratory marker for assessing myocardial damage in FD [45].

Finally, a recent study on DMD has indicated the matrix metalloproteinase-9 (MMP-9) as both a diagnostic and prognostic biomarker. Indeed, DMD patients showed a higher serum level of MMP-9 protein and tissue inhibitors of metalloproteinase-1 (TIMP-1) proteins with respect to controls, with MMP-9 levels being even higher in older, non-ambulant patients than in ambulant patients [46].

5.2. Predictive/therapeutic biomarker

Considering the heterogeneous nature of RDs, not all patients are expected to benefit from a newly available treatment. Hence the identification of a sub-group of patients likely to respond to a novel treatment is important both in terms of health, and in terms of cost-effectiveness [12]. To this end, a predictive, or therapeutic, marker must be able to discriminate between drug responders (patients gaining benefit from the therapy) and poor/low responders (Figure 4). Predictive biomarkers will therefore enable the most appropriate and efficacious treatments or interventions to be selected for each patient, thereby underpinning a personalised approach to treatment.

There are a few examples of therapeutic biomarkers useful in RDs, generally SNPs, as in typical pharmacogenetics, although some protein studies have also been reported. For instance, a pharmacological predictive biomarker has been reported in idiopathic nephrotic syndrome, a RD affecting the kidneys. Specifically, Wen et al. [47] found a significant difference in the serum proteome of steroid-sensitive nephrotic syndrome (SSNS) and steroid-resistant nephrotic syndrome (SRNS, OMIM: 256370) patients, predictive of their respective responses to treatment.

Another example of a predictive biomarker has been found for Gaucher disease (GD, OMIM: 230800), a rare recessive genetic disorder (approximate prevalence 1/100,000) caused by mutations in the GBA gene, which codes for a lysosomal enzyme, glucocerebrosidase. Although the clinical manifestations of this disease are extremely variable, ranging from non-neurological manifestations such as organomegaly, bone anomalies, and cytopenia to acute neurological forms, a recent time-course analysis of ferritin, chitotriosidase, haemoglobin, and platelets showed that the levels of these biomarkers undergo variation during the course of enzyme replacement therapy [48].

Sometimes the same biomarker can be useful in multiple scenarios and, for example TGF-c, in addition to serving a prognostic function in Marfan syndrome, could feasibly be used as a therapeutic biomarker in the same condition. Indeed, in a recent study, patients who respond-
ed to losartan used to reduce the aortic root dilatation rate, had higher baseline TGF-β levels but exhibited lower plasma TGF-β concentrations during losartan therapy [49].

Predictive biomarkers such as these are likely to play an increasingly important role in clinical practice, since evaluating the efficacy of a treatment/intervention is fundamental to making decisions about treatment choices, and therefore determining therapy outcomes.

6. Conclusions

Since the definition of biomarkers in 2001, their importance in clinical and research settings has increased dramatically due to their diagnostic/prognostic functions and their ability to monitor/predict disease stage, treatment response, and/or adverse effects. Indeed, the creation of an exhaustive catalogue of approved biomarkers may be the single most important innovation in healthcare, bringing considerable clinical and economic benefits. Although current research, both academic and corporate, is heavily focused on the development of drugs and companion diagnostic tests, in the future, biomarker discovery and development will be vital for tailoring medical care to individual patients. This will be especially important in the field of RDs, in which the discovery of efficacious biomarkers is likely to greatly facilitate the process of EMA approval and development of novel orphan drugs. In addition to being both and time and cost-effective, biomarker research also provides exciting opportunities to expand our knowledge of the physiopathological mechanisms behind rare and other diseases, helping to discriminate between distinct disease presentations and comorbidities, as well as predict the different impacts of concomitant medication, and various important demographic parameters such as gender, age, and ethnicity. In short, biomarker discovery represents a giant leap towards the ultimate goal of truly personalised medicine.

Acknowledgements

The BIO-NMD (http://www.bio-nmd.eu/) (European Union FP7 #241665) and Neuromics (European Union FP7 #305121) projects are acknowledged.

Author details

Chiara Scotton and Alessandra Ferlini*  
*Address all correspondence to: fla@unife.it  
Medical Genetics Unit, Department of Medical Sciences, University of Ferrara, Ferrara, Italy
References


[26] Harten IA, Zahr RS, Lemire JM, Machan JT, Moses MA, Doirion RJ, Curatolo AS, Rothman FG, Wight TN, Toole BP, Gordon LB. Age-dependent loss of MMP-3 in


