GENOMICS IN PATHOLOGY: THE HYPE, HOPE AND REALITY OF PRECISION MEDICINE

Prof Maritha Kotze

October 2017

GENOMICS IN PATHOLOGY: THE HYPE, HOPE AND REALITY OF PRECISION MEDICINE

Inaugural lecture delivered on 12 October 2017

Prof MJ Kotze
Division of Chemical Pathology
Faculty of Medicine and Health Sciences
Stellenbosch University

Printing: SUN MeDIA
ISBN: 978-0-7972-1708-9

ABOUT THE AUTHOR

Maritha Kotze was born in Bellville and schooled in the area. She obtained the degrees BSc (1980), BSc Honours (1982) and MSc (1984) *cum laude* at Stellenbosch University. She was employed as a cytogeneticist by the Provincial Administration of the Western Cape in 1981 and obtained her PhD on the molecular genetics of familial hypercholesterolaemia in 1990. In 1986, she received the AJ Brink Floating Trophy for the best presentation by a young scientist at the Academic Day of the Faculty of Medicine and Health Sciences at Stellenbosch University. Over the next decade she won seven publication prizes, including the Andries Blignaut Medal for the best article published in the *South African Medical Journal* in 1989. In recognition of her contribution to both the generation and the application of new knowledge in patient care, she was awarded the Rector's Award for Research Excellence in 1999. In 2000 she was appointed as the head of the Division of Human Genetics at the Stellenbosch University Faculty of Medicine and Health Sciences and promoted to associate Professor in 2001.

Her desire to translate research into clinical utility led to the co-founding of a molecular genetics company in 2002 with three of her previous PhD students. While working full-time at this laboratory over the next five years, she continued her academic research with doctoral supervision of both laboratory and clinical scientists. In 2004, she received a patent incentive award from the South African Medical Research Council (MRC) for research relating to familial hypercholesterolaemia. This innovation combined the analysis of distinct and shared genetic risk factors for cardiovascular disease and evolved into a pathology-supported genetic testing service from 2007. This involved development of the Gknowmix™ data-integration software program financed by the Support Program for Industry Innovation and an internship programme supported by the Technology Innovation Agency of South Africa.

Maritha has been registered at the Health Professions Council of South Africa as a medical scientist trained in Human Genetics since 1992, and obtained a second registration in the discipline Molecular Biology in 2013. Over the years she supervised more than 50 postgraduate students, including 15 PhDs. In her quest to make genomics applicable to clinical practice she registered an Applied Genetics Short Course for clinician education in 2014 and 2015. This initiative was extended as part of the P5 Africa Congress on Point of Care Testing and Personalised Medicine, which she co-organised in March 2016.

Her current research focus is in the field of cancer genetics and is supported by substantial research grants from the MRC and the Cancer Association of South Africa. This enabled her to advance the pathology-supported genetic testing concept from panel-based genetic testing to whole exome sequencing as the next frontier in personalised/precision medicine. She is currently employed in the Division of Chemical Pathology at the National Health Laboratory Service, Tygerberg Hospital, where she heads up the Personalised Medicine Research Group through a joint appointment at Stellenbosch University, from March 2016. She was rated an established researcher by the National Research Foundation, with 140 papers published in peer-reviewed journals, several invited book chapters and an H-index of 31.

ACKNOWLEDGEMENTS / DANKBETUIGINGS

I am grateful that both my home and my work environment are supportive and empowering. Throughout my life I received encouragement from my family, friends, colleagues and many research collaborators who believed in my ideas and aspirations. I am equally thankful for criticism, which shaped the way in which I reached the goals set for myself and my students.

Eerstens wil ek my man, Jan Daniël Kotze, ons twee seuns, Jan-Roux en Gerick, en my twee susters, Anita van Zyl en Annette Jansen van Vuuren, bedank vir hulle liefde en ondersteuning deur die jare. Jannie, sonder jou kookvernuf waarvoor jy bekendstaan in ons familie en onder jarelange vriende sou dit baie moeilik gewees het om tot hier in my loopbaan te vorder. Dankie vir die vaste beginsels wat jy in jou beroep as landdros uitgedra het en dat jou hande wat vir niks verkeerd staan nie ook ons huis in 'n veilige hawe omskep het. Ek dra hierdie intreerede op aan my liefdevolle moeder, Nettie (de Waal), en oorlede vader en broer met dieselfde naam, Archie Roux. Julle hardwerkendheid, moed en vaste beginsels is die rots waarop ek bou. Ek dink ook met deernis en baie liefde aan my oorlede skoonouers, wat 'n groot indruk op my lewe gemaak het weens hulle sterk geloofsoortuiging en omgee vir ander. Die mooi karaktertrekke wat albei ons ouers gekenmerk het, leef voort in Jan-Roux en Gerick. Hierdie twee jong manne is my inspirasie. Hulle maak my beker vol met hulle eie drome en talente, veral daar waar my vermoëns tekortskiet.

A special word of thanks to Profs Johann Schneider, Rajiv Erasmus, Annie Zemlin and Susan van Rensburg at the Department of Pathology, Stellenbosch University, as well as Nanette Spencer from the National Health Laboratory Service. Your understanding of how genomics could add value to pathology opened up a new approach to healthcare. Susan, I will always be grateful to you for coining the phrase 'pathology-supported genetic testing', when I was searching for a way to explain in simple terms that we do not apply genomics as a stand-alone laboratory science. A special word of thanks to Dr Lynn Horn, Dr Nicola Barsdorf, Francis Masiye and Professor Keymanthri Moodley for valuable advice related to various challenging aspects during the development of an integrated service and research platform. I gratefully acknowledge the contribution of Craig Watt, Lujane Nutt, Janine Cronjé, Lindiwe Whati, Petra Opperman, Theresa Brand, Elizna Kruger, Dirk de Klerk and Dieter Geiger who have steadfastly supported me on this journey.

Words cannot do justice to Professor Tony Bunn's efforts since the days that he headed up the Innovation Centre at the MRC. His energy, encouragement and enthusiasm were instrumental in the creation of the Gknowmix™ genetic knowledge integration system and his involvement remains crucial for success. My mentor from 2002, Emeritus Professor Peter Beighton of the University of Cape Town, is acknowledged for his vision as a clinical geneticist to support the use of genetic testing in risk management of complex disorders. Thank you for the honour to contribute a review article to your Festschrift published in the *South African Medical Journal* last year. A special word of thanks to our research team and all my students and intern medical scientists over the years who contributed to the development of the ideas and concepts presented in this review. I especially acknowledge Dr Karin Baatjes, who is currently registered as a PhD (Surgery) student, and our project manager, Dr Armand Peeters, my first doctoral student from Belgium many years ago. It is impossible to mention all my previous students and co-workers for their unique contributions, so I acknowledge the role of recent MMed and PhD graduates in the success of the MRC Strategic Health Innovation Partnership project: Dr Kathleen Grant, who lectures at the Cape Peninsula University of Technology; Dr Nicole van der Merwe, who is currently registered at the University of Cape Town for a master's degree in genetic counselling; and surgeons Drs Hein Pohl and Sheridan Santhia. I also thank their co-supervisors Profs Justus Apffelstaedt, Juanita Bezuidenhout and Colleen Wright, as well as Dr Rika Pienaar and Dr Ettienne Myburgh from the private sector who provided a real-world perspective from their experience in clinical practice. This enabled development of the testing algorithm that Professor Manie de Klerk used in 2009 to pioneer medical scheme reimbursement for certain multi-gene tests. Drs Henry Davis and Johann Raats are acknowledged for highlighting the need for an appropriate clinical protocol to facilitate the application of oncogenomics in South Africa.

Ek is baie dank verskuldig aan al die pasiënte wat oor die jare aan ons navorsingsprojekte deelgeneem het. Hul nalatenskap van data is geskenk tot voordeel van toekomstige generasies. Ek dink veral met deernis aan wyle Elna Symington, my vriendin uit skooldae wat ek onlangs weer op Facebook raakgeloop het. Hier het ek verneem van haar kankerdiagnose en kon haar laaste lewensjaar met haar deel. Haar erge newe-effekte met gebruik van effektiewe medikasie het dit vir my duidelik gemaak dat 'n vroeër fokus benodig word vir identifikasie van genetiese teikens vir kankerbehandeling en hoe dit lewenskwaliteit beïnvloed.

Last, but not least, I thank Profs Roes Pienaar and Dries Retief who introduced me to and fostered my enthusiasm for human genetics.

GENOMICS IN PATHOLOGY: THE HYPE, HOPE AND REALITY OF PRECISION MEDICINE

INTRODUCTION

The pathogenesis of a disease refers to the biological mechanisms underlying an acute, chronic or recurrent medical condition. The Greek words "pathos" (disease) and "genesis" (creation) were combined to reflect the key role of pathology in clinical diagnostics underpinned by genetic variation (Banay 1948). Genetics is the study of inheritance based on the structure and function of genes or chromosomal abnormalities, while genomics aims to identify the interrelationship between multiple genes and environmental risk factors to positively influence their combined effect on health and disease.

The impact of molecular pathology on health relies heavily on technological advances, which have evolved from invention of the light- and electron microscopes around the 20th century to 'game-changing' nucleic acid-based genomic innovations (Wall & Tonellato 2012). The terms 'precision medicine' and 'personalised medicine' are used interchangeably and refer to the tailoring of preventive and therapeutic interventions to the biological characteristics of the individual. Precision medicine should not be misinterpreted as the ability to design a unique treatment for every patient, but rather considered as incremental improvements to ultimately provide rapid, accurate and cost-effective genomic service delivery at the point of care. The promise to improve patient care at the bedside led to concerns that genomics is nothing more than hype and that the benefits are exaggerated. The reality is that cost barriers, misinterpretation of complex information and ethical concerns about genetic discrimination limit the application of precision medicine. These issues have intensified over the years as we moved from single- to multi-gene testing and next-generation sequencing (NGS) (Kotze 2016). NGS coupled with the use of sophisticated bioinformatics tools expanded the scope of genetic testing from a niche speciality for diagnosis of rare disorders and carrier screening to high-precision genome sequencing. Over the last three years third-generation nanopore sequencing became available on a portable device, with the promise to dramatically reduce costs and speed up clinical diagnoses (Loose 2017). Individualised treatment based on complex genomic information requires integration and interpretation of data beyond that to which practitioners are accustomed.

International standards for genetic testing were first defined by the World Health Organization (WHO) in 2004, the same year in which we published an article on the impact of genetics on insurance (Kotze et al. 2004a). The first test approved for clinical use under the WHO guidelines was for detection of the "Leiden" mutation in the *factor V* gene, named after the Dutch city where this gene variant associated with blood clotting was first identified (Bertina et al. 1994). Screening for the low-penetrance *factor V* Leiden (1691G>A) mutation forms part of a comprehensive cardiovascular disease (CVD) multi-gene assay described in the above-mentioned article (Kotze et al. 2004a), to help explain the difference between tests based on genetic technology and pathology tests that can also determine the likelihood of developing an inherited disorder. This assay, which includes genetic risk factors involved in cholesterol, folate, iron and drug metabolism, has recently been approved for development as a rapid point-of-care test kit under the South Africa-UK Newton Collaborative Research Development Programme in Precision Medicine. The aim is to incorporate genomics into a universally accepted body of knowledge routinely applied in clinical pathology, focused on both distinct genetic risk factors and shared disease pathways.

GENOMICS IN PATHOLOGY

In the pathology laboratory, body fluids, tissues and organs are examined to diagnose a disease as the basis for treatment. There are several ways in which the information contained in our genes can influence clinical decision making beyond that observed at the phenotypic level. Firstly, individual differences in pathology test results, such as high blood cholesterol or iron deficiency, may be caused by genetic variation that determines the severity of the condition. Linking a genetic component to biochemical abnormalities could explain hyper- or hypo-responsiveness to environmental stimuli. Regular follow-up may be required to monitor the effectiveness of therapeutic intervention based on the combination of the genetic and lifestyle risk factors identified. Secondly, medication side effects or failure may be explained or

prevented when the genetic make-up of a patient is taken into account. We studied these aspects in patients with depression and certain forms of cancer, which may show a strong correlation between host genetics and tumour histopathology. Thirdly, a family history of medical conditions associated with pathological features identified in a patient needs to be considered when genetic testing is performed to help distinguish between familial and lifestyle-related disease.

Genetic variants can be broadly categorised into those with high (>50%), moderate (20–50%) or low penetrance (10–20%) of clinical expression. In the case of breast cancer, a lifetime risk above 50% is sufficient to recommend family screening and prophylactic surgery in mutation-positive cases, even in the absence of other risk factors (Ghoussaini et al. 2013). Detection of less penetrant germline variants or tumour gene expression profiles associated with increased recurrence risk or metastatic potential, may identify the need for chemotherapy in addition to hormone treatment in early-stage breast cancer patients. Treatment based on variation in genes coding for drug-metabolising enzymes (pharmacogenomics) or those dependent on nutrient cofactors (nutrigenomics) for optimal enzyme activity, may apply to all three categories of risk. Van der Merwe et al. (2017) incorporated clinical data and laboratory test results obtained at the protein, RNA and DNA levels into a pathology-supported genetic testing (PSGT) algorithm. This approach is used to assist clinicians with the decision to pursue extended genetic testing beyond *BRCA1/2* mutation detection and to facilitate clinical interpretation of genome-scale sequencing data.

Table 1 explains how information contained in both the germline genome (inherited) and the tumour genome (somatic) can be applied in the risk management of breast cancer patients. In patients who experience treatment failure, moderate-risk cases, or where the boundaries between the above-mentioned risk categories are blurred, NGS using pre-selected gene panels or whole genome/exome sequencing (WGS/WES) may be applicable. We hope that the use of such an integrative model for application of genomics in pathology would facilitate improved patient management and disease prevention across the disease spectrum, ranging from monogenic disorders with a Mendelian inheritance pattern to complex multi-factorial diseases with a genetic component (Kotze et al. 2015).

Table 1. Application of PSGT in breast cancer risk management (modified from Kotze et al. 2015)

Process	Evaluation
Document clinical risk profile, relevant pathology data and family history in the context of genetic counselling support	Personal risk • Early age at diagnosis (<40 years) • Presence of bilateral breast cancer and/or ovarian cancer • Presence of co-morbidities and lifestyle risk factors • Early-stage, recurrence or metastatic breast cancer based on tumour characteristics
	Ancestral risk • A limited number of founder mutations may account for the disease in the majority of affected patients • This relates to high-risk populations (e.g. Ashkenazi Jewish, Afrikaners)
	Familial risk • Breast/ovarian cancer runs in the family from one generation to the next • Other cancers including colon, prostate and male breast cancer are important indicators, especially if at least four close relatives younger than 60 years are affected, three below 50 years or two below 60 years and ovarian cancer in the family

Process	Evaluation
	Tumour histopathology • Triple-negative breast cancer (oestrogen receptor- [ER-], progesterone receptor- [PR-], human epidermal growth factor receptor 2- [HER2-] negative) and patients with the overlapping basal-like subtype are at increased risk to carry a *BRCA1* mutation • Breast cancer in germline *TP53* mutation carriers is commonly HER2-positive • Germline mutations in *BRCA1* and *TP53* are predominantly associated with invasive ductal carcinoma, *BRCA2* with both ductal and lobular cancers and *CDH1* with invasive lobular carcinoma, but not with ductal carcinoma
	Treatment failure / side effects • CYP2D6 metabolises approximately 25% of commonly used prescription drugs and is therefore of relevance in a wide range of clinical domains • Use of certain antidepressants may influence tamoxifen response due to reduced CYP2D6 activity, found to be of particular concern in *BRCA1/2* mutation carriers
Define mutation screening approach prior to patient referral	First screen an affected family member • Aimed at detection of a pathogenic variant in *BRCA1/2* or other high-risk breast cancer gene • Implies high familial risk for recurrence or bilateral breast cancer and increased risk of ovarian/other cancer • Consider genetic testing for identification of treatment targets in affected patients based on histopathology
	Screen close relatives for family mutation • Exclude or confirm risk for inheritance of the family-specific mutation • Based on the outcome, the likelihood for development of cancer over the lifetime can be estimated in the context of treatment options
Pre-test genetic counselling in relation to familial risk identified	Test value in unaffected relative • A positive test would indicate a very high risk for cancer that requires intensified surveillance or prophylactic surgery with proven efficacy • A negative test for the family mutation may reduce the familial risk to that in the general population
	Test value in patient diagnosed with cancer • A positive test indicates a high risk for recurrence/bilateral breast cancer or ovarian/ other cancer • A negative test result for the family mutation means the cancer was caused by other risk factors not tested for (extended mutation screening may be recommended based on overall risk profile)
	Risk to offspring • A 50% chance with each pregnancy to pass on the faulty gene (e.g. *BRCA1/2* mutation-positive individuals) • No further risk implications for children when the causative family mutation was excluded

Process	Evaluation
Decide when to test in relation to familial risk	Above the age of 18 years • Women who need to have a bilateral oophorectomy usually prefer to know their risk before the operation • Women who consider breast surgery may use the information on mutation status as motivation for a bilateral mastectomy • Allows family planning and adequate screening to be timed between pregnancies • Decision about timing of risk-reduction surgery based on genetic risk and expression of disease occurring earlier in subsequent generations • Testing of germline or tumour DNA may be requested to predict therapeutic benefit of targeted therapies shown to be effective in genetic subgroups (e.g. PARP inhibitors in *BRCA1/2* mutation-positive cases) • Radiation therapy is also an important consideration due to risk associated with DNA double strand breaks associated with certain gene variants
Choose test option in relation to familial risk	Mutation-specific • Screen at-risk family members for a known *BRCA1/2* or other causative mutation previously identified in the index patient diagnosed with breast/ovarian cancer
	Population-specific • Screen for a limited number of mutations occurring at an increased frequency in cancer patients from certain population groups (e.g. Ashkenazi Jews, Xhosa, Afrikaners of European descent)
	Full gene screen • Screen for the cancer-causing mutation in breast cancer patients who tested negative during the initial screen for family/population-specific mutations or when these two test options are not indicated
Provide option for research participation in relation to storing, decoding or destroying of specimens used for routine genetic testing*	Research participation • *CYP2D6* pharmacogenomics testing as part of a chronic disease screening/monitoring programme aimed at lowering cumulative risk in breast cancer survivors • Ongoing studies in breast cancer survivors enable the research team to improve the tests and interpretation of results as new discoveries and actionable information become available in the scientific literature • This approach is of particular relevance for WGS/WES in non-responders/patients with severe drug side effects or genetically uncharacterised familial/early-onset cases
Clinical application	Data integration and report generation • Report contains genetic information related to the clinical condition and/or family history • Interpretation is focused on actionable genetic alterations validated in the laboratory • Genetic counselling and/or further expert clinical consultation or testing may be recommended based on the test outcome

*Routine genetic testing service delivery linked to the generation of a research database and biobank for improved quality assurance and validation of new tests against current standards

The implications of genetic test results need careful consideration in relation to treatment options and the potential impact on other family members prior to laboratory testing. Genetic counselling or family pedigree assessment will not always result in genetic testing. As genomic research gains momentum in Africa, the consent process could be enhanced by developing and testing innovative strategies to improve voluntary research participation and comprehension of goals for the implementation of precision medicine (Sawe et al. 2017a). Screening for germline mutations in high-penetrance genes provides information about the inheritance of cancer in a family and may simultaneously reveal biomarkers for targeted therapies, such as PARP inhibitors in patients with *BRCA1/2* mutations in their germline and/or tumour DNA. Therefore, the informed consent form used in research studies should encompass pre-test education about the risks, benefits and limitations of genetic testing. The use of stored patient information for validation studies performed at the interface between the laboratory and clinic could overcome the limitations of single health disciplines, provided that patient selection criteria are well defined and adhered to (Kotze et al. 2015).

INTEGRATING RESEARCH WITH SERVICE DELIVERY

Acceptance that genetic information is not sufficient to predict response to treatment led to the development of a new medical model for data acquisition, aimed at earlier adoption of pharmacogenomic applications in clinical practice (Baatjes et al. 2017). As explained on the informed consent form designed for 'application of personalized medicine using an integrated service and research approach' (Kotze 2016: S118), data can be extracted from our genomics database resource to 1) study the role of genetics in health and disease and 2) provide information that could help clinicians improve the medical treatment of their patients. Collective knowledge gained through comparative effectiveness studies involving advanced molecular technologies performed alongside standard pathology, has translated into immediate clinical benefits for many patients over recent years.

An important research area explored by our research group is to define whether genomic tests could add value to standard pathology. This was the focus of a doctoral study involving a 70-gene MammaPrint (Agendia BV, Amsterdam, Netherlands) microarray test that became available in South Africa after approval by the Food and Drug Administration in 2007. The MammaPrint pre-screen algorithm developed as a cost-saving strategy by Grant et al. (2013, 2015), takes advantage of the fact that the 70-gene profile excludes assessment of ER, PR and HER2 status already determined routinely in all newly-diagnosed breast cancers using less expensive immunohistochemistry (IHC) and fluorescence in situ hybridisation (FISH) tests. Pohl et al. (2016) demonstrated safe avoidance of chemotherapy in approximately 50% of the South African patient cohort studied. Clinical utility of MammaPrint was recently confirmed by level 1A evidence provided by the prospective MINDACT (Microarray in Node Negative and 1 to 3 Positive Lymph Node Disease May Avoid Chemotherapy) trial (Cardoso et al. 2016). As clearly demonstrated in this case, appropriate introduction of new companion diagnostics based on the biological characteristics of the tumour genome may outpace the reporting of randomised controlled trials that require lengthy follow-up for the final assessment of clinical outcomes. Although MammaPrint is one of the most expensive genomic tests available worldwide, our PSGT approach led to reimbursement by most medical schemes in South Africa and Namibia. Clinician education was provided through a series of Applied Genetics Workshops (2008-2015), aimed at the development of an ethics framework for integrating research with service delivery.

From 2010, an 80-gene microarray called BluePrint was added to the results of the MammaPrint test to 1) subdivide hormone-positive breast cancer into the luminal A and luminal B molecular subtypes, 2) identify endocrine treatment resistant hormone-positive tumours lacking ER-alpha function (basal-like), yet showing expression at the protein and single-gene RNA level, and 3) distinguish the HER2-enriched subtype as the predominant therapeutic target among HER2-positive breast cancers. Standard IHC/FISH methodologies to approximate these breast cancer subtypes according to the updated clinical guidelines, may lead to inaccurate results for selection of patients for anti-HER2 treatment (Press et al. 2016). The use of microarray technology, which captures the true biological profile of multiple genes regulating the functional activity of cancer pathways, was encouraged by Whitworth et al. (2014). These authors reported significantly different clinical outcomes compared with standard clinical molecular stratification.

Myburgh et al. (2016) highlighted the clinical overestimation of HER2 positivity in early-stage hormone-positive breast cancer in South African patients when relying only on standard pathology. The HER2-enriched subtype defined by the *ERBB2*,

GRB7, PERLD1 and *SYCP3* genes was identified in only approximately 20% of IHC/FISH HER2 positive cases. Previous studies indicated that omitting anti-HER2 therapy in patients with non-HER2 enriched tumours does not affect survival negatively. This finding is of particular relevance in resource-limited settings where significant disparity exists between patients who rely on state care and those who have access to private medical insurance. The incidence of HER2-positive breast cancer is estimated at 25% in southern Africa (Dickens et al. 2014), compared to 15 to 20% in most other countries. To improve access for all patients regardless of financial means, improved methods are needed to identify subsets of patients who may benefit most from biological therapies such as trastuzumab. In HER2-positive patients found to be resistant to this drug or others such as tamoxifen, targeted NGS of tumour-extracted DNA and/or germline WGS/WES may be applicable. In this context, Cabanillas et al. (2017) developed a combined molecular diagnostics platform for somatic and germline precision oncology. This application involves the analysis of 215 cancer genes: 97 for therapy selection and 148 for familial predisposition, of which 30 genes are informative for both applications. Realising this potential in clinical practice not only relies on successful classification of potentially causative genetic alterations into pathogenic or variants of uncertain clinical significance, but requires data integration that also takes the cumulative risk of intermediate phenotypes such as obesity and drug side effects into account.

An important focus for future research involves the genomic characterisation of high-risk African patient subgroups, such as the Kenyan breast cancer patients with aggressive, high-grade tumours described by Sawe et al. (2017b). The finding that immune cell infiltration correlates with proliferation highlights the potential of using immunotherapy strategies for treating these tumours. Although immunotherapy is not yet registered in African countries, treatment has been approved for at least eight types of cancer and 'hundreds of trials are pairing CTLA-4, PD-1 or PD-L1 inhibitors with chemotherapy, radiation and targeted therapies' (The Economist 2017; September 26: 3-4, 11). Although several South African cancer patients included in our database may benefit from immunotherapy, entering into a clinical trial has not been possible for most of them. Reprogramming the genome may also be achieved by 'chimeric antigen receptor' technology, but due to the uniqueness and complexity of individual immune responses, not all patients will respond as we may hope. When an actionable gene target or mutation is rare, patient recruitment into a trial is challenging. However, only a small number of participants may be required to reach statistical significance in clinical outcome studies.

Tumour heterogeneity and insufficient or lack of formalin fixed paraffin embedded (FFPE) tissue for NGS presents another challenge, which could be addressed by 'liquid biopsy' used as a surrogate for the entire tumour genome. This involves the sequencing of circulating cell-free DNA (cfDNA) shed into the patient's blood. As proof of principle, we recently identified the *TP53* c.818G>A (R273H) and *PTEN* c.950_953 del TACT (T319*fs) mutations in the tumour DNA of a South African patient with metastatic breast cancer where DNA extraction from solid tumour was not possible due to a lack of available FFPE tissue. Both mutations were excluded in germline DNA using conventional Sanger sequencing. TP53 protein accumulation and loss of PTEN is associated with resistance to commonly used anti-cancer drugs, as experienced in the South African patient. cfDNA now enables monitoring of cancer patients through identification of relevant biomarkers at various time-points (Heitzer et al. 2015). An example of the individualised decision making introduced by this technology can also be seen in patients with ovarian cancer. In a stroke of genetic irony, women carrying a mutation in the *BRCA 1 or 2* gene have a 40-80% lifetime risk of developing breast cancer, and a 20-40% lifetime risk for ovarian cancer. *BRCA1/2* mutation carriers, however, have a much better response to chemotherapy resulting in improved survival when compared to those with normal wild type *BRCA1/2* genes. Avastin (bevacizumab) is an anti-angiogenic drug designed to block a protein called vascular endothelial growth factor (VEGF). In cancer cells that produce too much VEGF, blocking this protein may prevent the formation of new blood vessels that supply tumour growth, but it increases the risk of arterial adverse events causing high-grade toxicity (Totzeck et al. 2017). VEGFR3 inhibition is associated with a 3- to 9-fold downregulation of *BRCA1* and *BRCA2* mRNA, thus potentially adding better treatment response from chemotherapy to those patients without a *BRCA* mutation (Lim et al. 2014).

The integrated research and service delivery platform was established to help us cope with the demand for clinical interpretation of ever-increasing genomic information, while striving to reduce costs and meet patient expectations (Kotze 2016).

UNCOILING DNA STRUCTURE AND FUNCTION

I was introduced to the application of genetics in patient care in 1981 when employed to count chromosomes under a microscope. We had to screen hundreds of cells for possible abnormalities in the structure or number of chromosomes. Cells in the human body contain 23 pairs of chromosomes that differ in length. We were trained to take photos of the best chromosomal arrangements, cut them out one by one, and stick each pair according to size and banding pattern on a piece of paper. Chromosomes 1 and 22 are the longest and shortest, respectively, and were grouped in pairs together with the sex-determining X- and Y-chromosomes as the 23rd pair. Each chromosome is made up of tightly coiled strands of DNA, consisting of four bases called adenine (A), thymine (T), guanine (G) and cytosine (C). The order of these bases determines the human DNA sequence, which is revealed in the uncoiled state of the DNA structure.

DNA sequencing was first performed in our laboratory in 1989, during my doctoral study on familial hypercholesterolaemia (FH). We used radioactive probe-labelling, conventional polymerase chain reaction (PCR) and toxic chemicals in polyacrylamide gels for electrophoresis to physically read each base in the gene region sequenced. The use of these basic methodologies led to the identification of three novel founder mutations (D154N, D206E and V408M) in the low-density lipoprotein receptor (LDLR) gene causing FH in the majority (~95%) of affected South African patients (Kotze et al. 1989; 1991; Leitersdorf et al. 1989). This explained why the Afrikaner population has the highest frequency of FH in the world, approximately five-fold higher than in the European population from which it originated (Seftel et al. 1980). This phenomenon, known as a founder effect, is characteristic of historically isolated populations descending from a small gene pool. It is also a major contributing factor to the high incidence of coronary heart disease in the Afrikaner population of South Africa.

Knowledge of the molecular basis of FH provided a definitive tool for presymptomatic diagnosis of FH in affected families, as evidenced by identification of the three Afrikaner founder mutations in 36% of individuals studied who did not fulfil the clinical criteria for FH (Kotze et al. 1992). In 1992 we identified both the D154N and the D206E mutations (compound heterozygote) in a 34-year old patient presenting at Tygerberg Hospital Lipid Clinic with a total cholesterol level of approximately 22 mmol/l. Three-weekly plasmapheresis during a 12-year period and cholesterol-lowering drug treatment led to a significant reduction in cholesterol, although never below 12 mmol/l (Kriek et al. 1992). Both mutations D154N and D206E detected in this patient were classified as receptor-defective with some residual protein function, compared to the more severe receptor-negative subtype of FH caused by mutation V408M. In a study performed in 35 homozygous FH patients mainly from the Johannesburg Lipid Clinic, reduction in low-density lipoprotein (LDL) cholesterol in five patients with the receptor-negative subtype was similar to that achieved in 30 patients with the receptor-defective FH subtype (Raal et al. 2000).

FH has a strong genetic basis with high LDL cholesterol levels present from birth. To our surprize we found that the same LDLR mutation is associated with variable clinical expression in the same family, ranging from occurrence of a myocardial infarction from an early age (<50 years) to good health into advanced age (>80 years), despite high cholesterol levels in both cases (Kotze et al. 1993a). The type or severity of the gene defect furthermore affects serum cholesterol levels differentially, although it was not sufficient to explain clinical variability in coronary events (Kotze et al. 1993b). These findings prompted the development of a molecular diagnostic service for FH in South Africa to facilitate optimal cholesterol-lowering treatment and genetic counselling of affected families (Kotze et al. 1994). A non-radioactive multiplex PCR assay was also developed to improve cost-effectiveness of FH testing in clinical practice (Kotze et al. 1995). This method was applied in different laboratories in South Africa based on our publication, and was only replaced 10 years later by a reverse-hybridisation strip-assay covering a larger number of mutations identified in the diverse South African population (Kotze et al. 2003; Kotze & Thiart 2003).

In the following years we performed many collaborative studies to determine the cause or consequences of FH in patients of different ancestries (Callis et al. 1998), including Indian (Kotze et al. 1997; Langenhoven et al. 1996), Black (Peeters et al. 1998; Thiart et al. 2000) and Coloured (Loubser et al. 1999). Although most mutations have been identified in the coding region of the LDLR gene, we demonstrated the functional significance of promoter mutations (-59, -92) in transfected cells (Peeters et al. 1998; Scholtz et al. 1999) and detected both novel and known gene variants at exon-intron splice junctions (Peeters et al. 1999). Mutations in the apolipoprotein B-100 gene also found in patients presenting with clinical features of FH are

rare in the South African population (Rubinsztein et al. 1995). Genetic variation in other genes studied in FH patients could explain variation in triglycerides and high-density lipoprotein (HDL) cholesterol levels (Ehrenborg et al. 1997; Pimstone et al. 1995). Determination of lipoprotein(a) levels in FH patients using sib pair analysis, showed an additive effect of a minor allele in the presence of a major allele (Lingenhel et al. 1998). In a study performed in 221 South African children from 85 FH families, plasma cholesterol levels overlapped considerably between mutation-positive and -negative cases, suggesting that modifiable lifestyle factors remain important in children with heterozygous FH (Kotze et al. 1998). Many genetic and environmental factors influence the risk of coronary heart disease in FH patients, but are not sufficient on their own to cause a heart attack. Exclusion of founder mutations in clinically diagnosed FH patients led to extended mutation screening, which resulted in the addition of approximately 50 different LDLR gene mutations to electronic genetic databases (Varret et al. 1998, Villeger et al. 2002), as well as the identification of a new locus for FH (Varret et al. 1999).

Our experience from extensive studies performed in families with FH provided a paradigm for the future (Kotze & Callis 1999). In 2001, when the first draft of the human DNA sequence was published (Lander et al. 2001), we reported the first prenatal diagnosis of FH in the high-risk South African population (Vergotine et al. 2001a). This important milestone demonstrated the value of a population-directed screening strategy to identify FH patients in populations with an enrichment of specific LDLR gene mutations. In this family study, mutation D206E previously identified in germline DNA of both the mother and the father (FH heterozygotes) as well as their two children (one homozygote) was not detected in DNA directly amplified from the mother's amniotic fluid. Cells were also cultured *in vitro* for a second round of genetic testing to exclude the possibility of a false-negative result in the foetus. This was done due to the small number of cells in the amniotic fluid sample. At birth we collected cord blood for mutation analysis again to exclude maternal contamination. This confirmed the previous result of the FH test.

Knowledge of the DNA structure of the LDLR gene made early diagnosis of FH, including prenatal diagnosis, a reality. However, despite extensive awareness and publicising of hypercholesterolaemia as a significant risk factor for coronary heart disease, less than 10% of heterozygous FH cases in South Africa are correctly diagnosed and treated optimally (Marais et al. 2004). Vergotine et al. (2001b) demonstrated that determination of total cholesterol levels in FH families with known mutations fails to provide the correct diagnosis in nearly 30% of patients when the 95th percentile for age and gender is used, and in 12% of cases when the 80th percentile is used. Of the estimated 60 to 70% of South Africans with elevated cholesterol levels, 5 to 10% will have FH. This distinction is very important for treatment considerations, as high-risk FH patients require long-term drug treatment in addition to nutrition and lifestyle intervention as the treatment of choice in most patients with less severe forms of dyslipidaemia. Male FH patients have a more than 50% risk of coronary heart disease by age 50 and in women the risk is approximately 30% by age 60. The average age of death is 45 years in men with FH.

Apart from the *LDLR* and *APOB* genes, genome-wide association studies (GWAS) and NGS enabled identification of at least four more major FH genes *(PCSK9, LDLRAP1, LIPA and STAP1)*, as well as a polygenic form of FH. Talmud et al. (2013) developed a LDL cholesterol genetic score to help distinguish between patients with polygenic and monogenic hypercholesterolaemia. In four South African individuals from the same family selected for WES based on exclusion of the three Afrikaner founder mutations, none of the above-mentioned genes could explain their abnormal cholesterol levels. Detection of the relatively common ε4 allele of the apolipoprotein E *(APOE)* ε2/3/4 polymorphism that forms part of the genetic risk score, was however detected in the father and sister of the index case. This gene variant, which occurs in 30-40% of the general population, has previously been identified as an important biomarker for differential diagnosis of monogenic versus polygenic FH in the South African population (Kotze et al. 1993b; Lückhoff et al. 2015). The fact that *APOE* provides a genetic link between Alzheimer's disease and CVD is considered clinically useful in relation to both the positive and negative genotyping results obtained (Kotze & Van Rensburg 2012).

SHORTFALL IN HUMAN DNA SEQUENCING DATA

Insights gleaned from the entire sequence of the human genome first published in 2001 introduced a new era in medicine with profound implications for health and disease (Lander et al. 2001). Approximately 20 000 genes were identified (Ezkurdia et al. 2014), which are far less than the estimated 40 000 to 100 000 two decades earlier (Fields et al. 1994). This surprisingly low number for our species, which is no more than the

chromosome content of nematode worms, was difficult to contemplate given the vast difference in complexity of the two. This meant that the one-disease one-gene relationship forming the basis of human genetics in the past is an exception rather than the rule.

Explanations for the relatively small number of human genes include the existence of shared disease pathways and epigenetics involving gene-environment interaction. These networks are influenced by the genetic backgrounds of patients through effects that may be too small to be detected by usual statistical means. GWAS resulted in the identification of thousands of single nucleotide polymorphisms (SNPs) implicated as risk modifiers in both monogenic disorders and complex disease traits. Diseases of complex origin are characterised by multiple intermediate phenotypes, with a component of quantitative genetics involved in their pathogenesis (Blanco-Gomez et al. 2016). GWAS failed to explain the high level of discrepancy found between the proportion of phenotypic variance expected to be caused by genetic influences and that actually accounted for by sequence variants. This so-called missing heritability of the human genome (Manolio et al. 2009) has also been investigated by advanced NGS techniques that became available over the last decade. Genome-scale NGS enables the identification of rare variants not readily detectable by GWAS. However, neither the *common disease / common variant* hypothesis investigated by GWAS nor the *common disease / rare variant* hypothesis investigated by NGS were able to sufficiently explain the 'missing heritability' of the human genome.

Heritability is estimated through phenotypic correlations between family members and twin studies. The availability of advanced NGS technologies caused a shift from the focus on genes mutated in monogenic to complex disorders, which result from effects of multiple genes and environmental exposures. Genetic studies of intelligence, for example, confirmed a polygenic nature involving more than 50 genes (Sniekers et al. 2017). NGS was also used to perform a transcriptomic study in children with Down syndrome, the most frequent and recognisable cause of intellectual disabilities. Eighty genes were selected for further analysis, of which only two (*HLA DQA1* and *HLA DRB1*) were significantly down-regulated among patients who scored the lowest for intelligence (Mégarbané et al. 2013). Down syndrome is caused by an extra chromosome 21 and presents as the most common infant genetic abnormality detectable under a microscope. About half of all affected children are born with a heart defect and gradual cognitive decline may be experienced with aging, resulting in Alzheimer's disease in 50 to 70% of adults in their fifties or sixties. The increased prevalence of Alzheimer's disease in Down syndrome was partly explained by excessive genetic material of the amyloid beta precursor protein gene on chromosome 21, which would not be visible through a microscope (Klosowska 2017). The ε4 allele of the *APOE* gene on chromosome 19 also plays a role; it not only contributes to Alzheimer's disease risk, but also increases maternal meiosis segregation II errors in Down syndrome (Avramopoulos et al. 1996; Bhaumik et al. 2017).

APOE ε2/3/4 is the most studied polymorphism displaying pleiotropic effects in humans, first described in 1980 as a biochemical abnormality of lipid metabolism (Utermann et al. 1980). However, translation of *APOE* genotyping results into treatment designs remains a challenge, as the majority (75%) of individuals with one copy (heterozygous) of the *APOE* ε4 allele do not develop Alzheimer's disease during their lifetime, and half of people with Alzheimer's disease do not carry this risk-associated allele (Farrer et al. 1997). Although the clinical utility of *APOE* genotyping is insufficient for diagnosis of Alzheimer's disease, detection of the risk allele emphasises the importance of a healthy lifestyle and may raise the quality of care towards Alzheimer's disease prevention (Villeneuve et al. 2014). Studies performed in South African patients demonstrated an apparent epistatic effect of the *APOE* ε4 allele on age of disease onset in patients with dementia (Van Rensburg et al. 2000), as well as significant effects on lipid levels in the general population (Kotze et al. 1993b). *APOE* genotyping performed in more than 500 South African individuals participating in a chronic disease screening programme confirmed the beneficial effect of physical activity in relation to genetic risk factors and dyslipidaemia (Lückhoff et al. 2015). The genetics of Alzheimer's disease, similar to most other complex diseases, forms part of a continuum with highly penetrant, early-onset familial disease at one end of the spectrum and combinations of polymorphic variants with mild effects contributing to sporadic late-onset Alzheimer's disease towards the other end (Van Cauwenberghe et al. 2016). With the availability of NGS the tools are now available to determine the genetic footprint associated with a low-, intermediate- and high-risk clinical profile for AD linked to more than 20 genomic loci (De Roeck et al. 2017).

Future prospects of combining long-read (>150 kb) third- and *in situ* fourth-generation sequencing with well-established NGS technologies suitable for WES/

WGS, has generated much excitement in the genomics community (Jain et al. 2016; McGinn et al. 2016). Magi et al. (2016) reported that the MinION nanopore sequencing device can be readily used to detect genomic regions involved in copy number variants with high accuracy, outperforming other state-of-the-art methods in terms of both sensitivity and specificity. Applications include analysis of infectious agents as demonstrated on-site in West Africa with the outbreak of the Ebola virus, and aneuploidy detection in prenatal samples shown to reduce the analysis from 1-3 weeks to less than 4 hours in a clinical setting. Long reads from previously 'unsequenceable' genomic regions, coupled with the use of sophisticated bioinformatics tools (Chu et al. 2017; Liu et al. 2017; Yang et al. 2017), may provide genomic and epigenomic tumour and germline diagnoses within the same day. Due to the low cost of entry and portability, the MinION sequencer furthermore provides an excellent tool for teaching of students performing their own NGS experiments.

NATURE MEETS NURTURE

The nature (genetics) versus nurture (environment) debate has been brewing for as long as I can remember. For geneticists trained in the pre-genomic era the phenomenon of 'missing heritability' is problematic. This term refers to the fact that genetics cannot fully explain clinical variability as the number of ways in which genes could combine to produce different phenotypes is almost infinite. Blanco-Gomez et al. (2016) enlightened the problem of missing heritability in complex human diseases with a strong case for the role of intermediate phenotypes that can be measured by standard pathology tests and clinical assessments. These include CVD risk factors such as hypercholesterolaemia and obesity, each with its own set of cumulating genetic and environmental risk factors. Studies performed by my research group over the past three decades demonstrated the value of combining distinct (founder mutations) and shared (pleiotropic) genetic risk factors into the above-mentioned PSGT algorithm. Common CVD risk factors not only contribute to the pathogenesis of the FH phenotype or breast cancer as examples of main complex diseases, but are themselves influenced by third-order phenotypes in relation to these conditions. Variations in obesity-related genes may cause a tendency for chronic, low-grade inflammation that increase the risk of the metabolic syndrome. This condition is characterised by the presence of at least three intermediate phenotypes, including central obesity, dyslipidaemia (low HDL-cholesterol, high triglycerides), hypertension, and/or

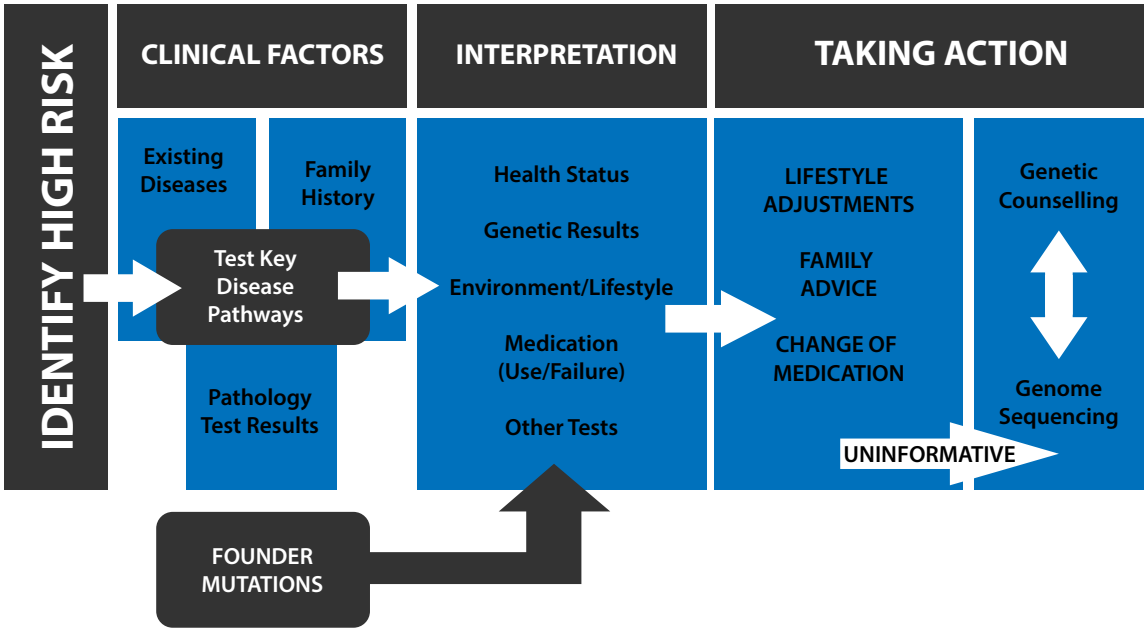

Figure 1. Testing algorithm using PSGT as a pre-screen step considering both pleiotropic and founder effects to determine eligibility for NGS. Focus areas for analysis of shared disease pathways relate to features of the metabolic syndrome, considered a unifying risk factor for many chronic diseases, which involves abnormalities in lipid and lipoprotein metabolism, DNA methylation and mismatch repair, haemostasis and inherited thrombophilia, haem synthesis and iron homeostasis, as well as drug metabolism (Kotze 2016).

glucose intolerance. Each of these metabolic syndrome components, routinely assessed in medical scheme wellness programmes, contributes its own combination of genetic and lifestyle risk factors to the overall risk profile of an individual. Figure 1 illustrates the translation of this knowledge into the PSGT algorithm, developed to facilitate the move from single- to multi-gene testing and WGS/WES (Kotze 2016). This approach requires pre-selection of patients and data integration for clinical decision making.

The genetic diversity of the South African population provides a valuable resource for biomarker discovery. The identification of founder mutations characteristic of the homogeneous Afrikaner gene pool led to the development of cost-effective assays for the accurate diagnosis of relatively common inherited disorders. These include the above-mentioned autosomal dominant form of FH, variegate porphyria (VP), and hereditary haemochromatosis (HH), a preventable iron overload disorder. Since discovery of the most common *HFE* gene mutations (C282Y and H63D) in 1996, rapid clinical benefit was seen through drastic reduction in liver biopsies (Feder et al. 1996). The importance of integrating genomics with pathology is clearly demonstrated by the variable clinical expression of haemochromatosis (Table 2). Iron uptake, storage and release are regulated at the gene level, with at least six genes *(HFE, HJV, HAMP, TFR2, SLC40A1, FTL)* implicated in oxidative stress, inflammation and ultimately loss of iron homeostasis resulting in clinical manifestation of HH. The high prevalence of the founder-related *HFE* C282Y and H63D mutations associated with these pathogenic events facilitated an improved diagnostic service for HH in South Africa (De Villiers et al. 1999; Kotze et al. 2004b) and elsewhere.

Genetic variation in iron-related genes such as *HFE* exerts a direct pleiotropic effect on iron parameters, which is partly regulated by hepcidin at the nexus of infection and anaemia. Hepcidin is upregulated by inflammation and reduced expression of the transmembrane serine protease S6 (TMPRSS6). Variation in this gene is associated with retention of iron in macrophages and decreased iron absorption (Traglia et al. 2011). Particular care should be taken to prevent misdiagnosis of HH in patients with hyperferritinaemia caused by the insulin resistance hepatic iron overload syndrome, also known as dysmetabolic iron overload. The connection of these diverse medical conditions through dysfunction of the iron metabolism pathway supports the implementation of preventative and therapeutic strategies based on the understanding of the mechanisms through which genetic and environmental factors contribute to disease development, progression and recurrence (Kotze et al. 2009). Blood donors

Table 2. Symptoms and signs of inherited iron overload associated with a pleiotropic effect (Kotze et al. 2009)

Symptoms and signs	Medical conditions
Abnormal liver function	Arrhythmias
Abdominal pain (unexplained)	Arthritis, arthralgia
Bronzing of the skin	Cardiomyopathy
Amenorrhoea (no menstrual periods, women)	Chronic fatigue
Anterior pituitary failure	Chronic liver disease
Frequent diarrhoea	Cirrhosis of the liver
Hyperferritinaemia	Depression
Impotence (men)	Diabetes mellitus Type 1
Insulin resistance	Diabetes mellitus Type 2
Joint pain	Fatty liver disease
Loss of body hair	Hepatocellular carcinoma
Loss of libido	Infertility
Mood swings	Metabolic syndrome
Muscle pain	Porphyria cutanea tarda
Skin pigmentation	Testicular atrophy (men)
Weakness	

and vegetarians with the HH genotype are unlikely to develop clinically manifested HH due to protection against iron overload. Lack of clinical manifestation of the HH genotype has also been described in a subgroup of South African patients diagnosed with multiple sclerosis (Kotze et al. 2006). This finding may partly be explained by interaction between the *HFE* and *TMPRSS6* genes (Van Rensburg et al. 2012). Based on this knowledge, two children with multiple sclerosis were successfully treated with iron supplementation at Tygerberg Hospital (Van Toorn et al. 2010; Van Rensburg et al. 2015). Iron is a vital requirement for cholesterol and lipid biosynthesis, which are both key components of myelin known to be negatively affected by iron deficiency.

According to Beutler (2007) there are no single gene diseases. Whether we are dealing with an autosomal dominant form of FH, autosomal recessive form of HH, or subtypes of multi-factorial diseases such as MS, 'individuals with the same genotype have very different clinical phenotypes' (Beutler 2007: 5).

CONCLUSION

The tools developed for application of precision medicine has moved out of the laboratory into clinical practice. New knowledge generated through genomic research will continue to deliver effective drugs and bring hope for a cure of the treated condition. An in-depth understanding of genomics and how complex nature is helps us to separate hype from reality. The application of advanced molecular technologies comes with the responsibility to act on the information they provide and, most importantly, the need for a system to manage patient expectations before embarking on genome-scale analyses (Kotze et al. 2013, van der Merwe et al. 2017). Without a well-defined strategy to overcome the limitations of genetics alone to account for clinical manifestation of complex diseases and response to treatment, precision medicine will be "caught between hype, hope and the real world" (Dammann & Weber 2012: 91).

A major challenge for translation of genetic research into clinical practice is the shortage of clinicians educated in the field of genomics (Aspinall & Hamermesh 2007). Advanced NGS technologies has broken the 'sound barrier' of human genetics through the generation of large volumes of sequence data at a speed higher than ever imagined possible (Bahassi el & Stambrook 2014). While students today enter medical science at this mind-boggling level, I was fortunate to live through the process. As part of a team, we could tackle and overcome each new barrier from a multi-disciplinary perspective. I came to understand why patients with the same genetic disorder or cancer subtype enjoyed healthy aging, while others who received the same medical treatment, died prematurely. Sometimes not because of the failure of gene-targeted therapies, but due to their toxic effects that frequently include adverse CVD events. Information gaps are steadily filled by ongoing data collection and health outcome studies linked to routine service delivery, towards the right treatment, for the right patient, at the right time.

'It is awe-inspiring, to realize that we have caught the first glimpse of our own instruction book, previously known only to God' - Dr Francis Collins.

REFERENCES

Aspinall MG, Hamermesh RG. Realizing the promise of personalized medicine. Harv Bus Rev 2007; 85(10): 108–117.

Avramopoulos D, Mikkelsen M, Vassilopoulos D, et al. Apolipoprotein E allele distribution in parents of Down's syndrome children. Lancet 1996; 347: 862–865.

Baatjes KJ, Conradie M, Apffelstaedt JP, Kotze MJ. Pharmacogenetics of aromatase inhibitors in endocrine responsive breast cancer: lessons learnt from tamoxifen and CYP2D6 genotyping. Anticancer Agents Med Chem 2017; Apr 12. [Epub ahead of print]

Bahassi el M, Stambrook PJ. Next-generation sequencing technologies: breaking the sound barrier of human genetics. Mutagenesis 2014; 29(5): 303–310.

Banay GL. An introduction to medical terminology I. Greek and Latin derivations. Bull Med Libr Assoc 1948; 36(1): 1–27.

Bertina RM, Koeleman BP, Koster T, et al. Mutation in blood coagulation factor V associated with resistance to activated protein C. Nature 1994; 369(6475): 64–67.

Beutler E. Iron storage disease: Facts, fiction and progress. Blood Cells Mol Dis 2007; 39(2): 140–147.

Bhaumik P, Ghosh P, Ghosh S, et al. Combined association of presenilin-1 and apolipoprotein E polymorphisms with maternal meiosis II error in Down syndrome births. Genet Mol Biol 2017; 40(3): 577-585.

Blanco-Gomez A, Castillo-Lluva S, Del Mar Saez-Freire M, et al. Missing heritability of complex diseases: enlightenment by genetic variants from intermediate phenotypes. Bioessays 2016: 38: 664–673.

Cabanillas R, Diñeiro M, Castillo D, et al. A novel molecular diagnostics platform for somatic and germline precision oncology. Mol Genet Genomic Med 2017; 5(4): 336-359.

Callis M, Jansen S, Thiart R, et al. Mutation analysis in familial hypercholesterolemia patients of different ancestries: identification of three novel LDLR gene mutations. Mol Cell Probes 1998; 12: 149–152.

Cardoso F, Van't Veer LJ, Bogaerts J, et al. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. N Engl J Med 2016; 375: 717–729.

Chu J, Mohamadi H, Warren RL, et al. Innovations and challenges in detecting long read overlaps: an evaluation of the state-of-the-art. Bioinformatics 2017; 33(8): 1261-1270.

Dammann M, Weber F. Personalized medicine: caught between hope, hype and the real world. Clinics 2012; 67(Suppl 1): 91–97.

De Roeck A, Van den Bossche T, van der Zee J, et al. Deleterious ABCA7 mutations and transcript rescue mechanisms in early onset Alzheimer's disease. Acta Neuropathol 2017, 134(3): 475-487.

De Villiers JNP, Hillermann R, De Jong G, et al. High prevalence of the Cys282Tyr HFE mutation facilitates an improved diagnostic service for hereditary haemochromatosis in South Africa. S Afr Med J 1999; 89: 279–282.

Dickens C, Duarte R, Zietsman A, et al. Racial comparison of receptor-defined breast cancer in southern African women: subtype prevalence and age-incidence analysis of nationwide cancer registry data. Cancer Epidemiol Biomarkers Prev 2014; 23: 2311–2321.

Ehrenborg E, Clee SM, Pimstone SN, et al. Ethnic variation and in vivo effects of the -93 promoter variant in the lipoprotein lipase gene. Arterioscler Thromb Vasc Biol 1997; 17: 2672–2678.

Ezkurdia I, Juan D, Rodriguez JM, et al. Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. Hum Mol Genet 2014; 23(22): 5866–5878.

Farrer LA, Cupples LA, Haines JL, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E and Alzheimer disease: a meta-analysis. APOE and Alzheimer disease Meta-Analysis Consortium. JAMA 1997; 278: 1349–1356.

Feder JN, Gnirke A, Thomas W, et al. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. Nat Genet 1996; 13(4): 399–408.

Fields C, Adams MD, White O, Venter JC. How many genes in the human genome? Nat Genet 1994; 7(3): 345–346.

Ghoussaini M, Pharoah PDP, Easton DF. Inherited genetic susceptibility to breast cancer: the beginning of the end or the end of the beginning? Am J Pathol 2013; 183: 1038–1051.

Grant KA, Apffelstaedt JP, Wright C, et al. MammaPrint Prescreen Algorithm (MPA) reduces chemotherapy in patients with early stage breast cancer. S Afr Med J 2013; 103(8): 522–526.

Grant KA, Pienaar FM, Brundyn K, et al. Incorporating microarray assessment of HER2 status in clinical practice supports individualised therapy in early-stage breast cancer. Breast 2015; 24(2): 137–142.

Heitzer E, Ulz P, Geigl JB. Circulating tumor DNA as a liquid biopsy for cancer. Clin Chem 2015; 61(1):112-123.

Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol 2016;17(239): 1-11.

Klosowska A, Cwiklinska A, Kuchta A. Down syndrome, increased risk of dementia and lipid disturbances. Dev Period Med 2017; 21(1): 69–73.

Kotze MJ, Langenhoven E, Warnich L, et al. The identification of two low-density lipoprotein receptor gene mutations in South African familial hypercholesterolaemia. S Afr Med J 1989; 76: 399–401.

Kotze MJ, Langenhoven E, Warnich L, et al. The molecular basis and diagnosis of familial hypercholesterolaemia in South African Afrikaners. Ann Hum Genet 1991; 55: 115–121.

Kotze MJ, Langenhoven E, Kriek JA, et al. DNA screening of hyperlipidemic Afrikaners for familial hypercholesterolemia. Clin Genet 1992; 42: 43–46.

Kotze MJ, Davis HJ, Bissbort S, et al. Intrafamilial variability in the clinical expression of familial hypercholesterolemia: importance of risk factor determination for genetic counselling. Clin Genet 1993a; 43: 295–299.

Kotze MJ, De Villiers WJS, Steyn K, et al. Phenotypic variation among familial hypercholesterolemics heterozygous for either one of two Afrikaner founder LDL receptor mutations. Arterioscler Thromb 1993b; 13: 1460–1468.

Kotze MJ, Langenhoven E, Theart L, et al. Report on a molecular diagnostic service for familial hypercholesterolemia in Afrikaners. Genet Counseling 1994; 5: 15–21.

Kotze MJ, Theart L, Callis M, et al. Nonradioactive multiplex PCR screening strategy for the simultaneous detection of multiple low density lipoprotein receptor gene mutations. PCR Methods Applic 1995; 4: 352–356.

Kotze MJ, Loubser O, Thiart R, et al. CpG hotspot mutations at the LDL receptor locus are a frequent cause of familial hypercholesterolemia among South African Indians. Clin Genet 1997; 51: 394–398.

Kotze MJ, Peeters AV, Loubser O, et al. Familial hypercholesterolemia: potential diagnostic value of mutation screening in a pediatric population of South Africa. Clin Genet 1998; 54: 74–78.

Kotze MJ, Callis M. MED-PED FH: A paradigm for other common monogenic diseases in South Africa. Atherosclerosis 1999; 144 (Suppl): 467–468.

Kotze MJ, Kriegshäuser G, Thiart R, et al. Simultaneous detection of multiple familial hypercholesterolaemia mutations facilitates an improved diagnostic service in South African patients at high risk of cardiovascular disease. Mol Diagn 2003; 7: 169–174.

Kotze MJ, Thiart R. Genetics of dyslipidaemia. CME Journal 2003; 21: 399–402.

Kotze MJ, Schorn D, Coetzer P. The impact of genetic testing on life insurance. J Genomics Afr Soc 2004a; 1(1): 1–11.

Kotze MJ, De Villiers JNP, Bouwens CSH, et al. Molecular diagnosis of hereditary haemochromatosis: application of a newly-developed reverse-hybridisation assay in the South African population. Clin Genet 2004b; 65: 317–321.

Kotze MJ, De Villiers JN, Warnich L, et al. Lack of clinical manifestation of hereditary haemochromatosis in South African patients with multiple sclerosis. Metab Brain Dis 2006; 21: 109–120.

Kotze MJ, Van Velden D, Van Rensburg SJ, Erasmus R. Pathogenic mechanisms underlying iron deficiency and iron overload: new insights for clinical application. J Int Fed Clin Chem Lab Med 2009: 1–15.

Kotze MJ, Van Rensburg SJ. Pathology supported genetic testing and treatment of cardiovascular disease in middle age for prevention of Alzheimer's disease. Metab Brain Dis 2012; 27: 255–266.

Kotze MJ, Van Velden DP, Botha K, et al. Pathology-supported genetic testing directed at shared disease pathways for optimized health in later life. Personalized Med 2013; 10(5): 497–507.

Kotze MJ, Lückhoff HK, Peeters AV, et al. Genomic medicine and risk prediction across the disease spectrum. Crit Rev Clin Lab Sci 2015; 19: 1–15.

Kotze MJ. Application of advanced molecular technology in the diagnosis and application of genetic disorders in South Africa. S Afr Med J 2016; 106: S114–S118.

Kriek JA, Steyn K, Weight MJ, et al. Response to treatment with simvastatin in a compound FH heterozygote. Atherosclerosis 1992; 93: 165–167.

Magi A, Giusti B, Tattini L. Characterization of MinION nanopore data for resequencing analyses. Brief Bioinform 2016, Aug 24 [Epub ahead of print].

Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409(6822): 860–921.

Langenhoven E, Warnich L, Thiart R, et al. Two novel point mutations causing receptor-negative familial hypercholesterolemia in a South African Indian homozygote. Atherosclerosis 1996; 125: 111–119.

Leitersdorf E, Van der Westhuyzen DR, Coetzee GA, Hobbs HH. Two common low density lipoprotein receptor gene mutations cause familial hypercholesterolemia in Afrikaners. J Clin Invest 1989; 84(3): 954–961.

Lim J, Yang K, Taylor-Harding B. VEGFR3 Inhibition chemosensitizes ovarian cancer stemlike cells through down-regulation of BRCA1 and BRCA2. Neoplasia 2014; 16(4): 343–353.

Lingenhel A, Kraft H-G, Kotze MJ, et al. Concentrations of the atherogenic Lp(a) are elevated in FH. Eur J Hum Genet 1998; 6: 50–60.

Liu Q, Zhang P, Wang D, et al. Interrogating the „unsequenceable" genomic trinucleotide repeat disorders by long-read sequencing. Genome Med 2017; 9(1): 65. doi: 10.1186/s13073-017-0456-7.

Loose MW. The potential impact of nanopore sequencing on human genetics. Hum Mol Genet 2017; July 17 [Epub ahead of print].

Loubser O, Marais AD, Kotze MJ, et al. Founder mutations in the LDL receptor gene contribute significantly to the familial hypercholesterolemia phenotype in the indigenous South African population of mixed ancestry. Clin Genet 1999; 55: 340–345.

Lückhoff HK, Brand T, Van Velden DP, et al. Clinical relevance of apolipoprotein E genotyping based on a family history of Alzheimer's disease. Curr Alzheimer Res 2015; 12(3): 210–217.

Magi A, Giusti B, Tattini L. Characterization of MinION nanopore data for resequencing analyses. Brief Bioinform. 2016, Aug 24 [Epub ahead of print].

Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009; 461(7265): 747–753.

Marais AD, Firth JC, Blom DJ. Familial hypercholesterolemia in South Africa. Sem Vasc Med 2004; 4: 93–95.

McGinn S, Bauer D, Brefort T et al. New technologies for DNA analysis--a review of the READNA Project. N Biotechnol 2016; 33(3): 311-330.

Mégarbané A, Noguier F, Stora S. The intellectual disability of trisomy 21: differences in gene expression in a case series of patients with lower and higher IQ. Eur J Hum Genet 2013; 21(11): 1253–1259.

Myburgh EJ, Langenhoven L, Grant KA, et al. Clinical overestimation of HER2 positivity in early estrogen and progesterone receptor–positive breast cancer and the value of molecular subtyping using BluePrint. J Global Oncol 2016; 3: 314–322.

Peeters AV, Kotze MJ, Scholtz CL, et al. The 3-basepair deletion in repeat 1 of the LDL receptor promoter reduces transcriptional activity in a South African Pedi. J Lipid Res 1998; 39: 1021–1024.

Peeters AV, Thiart R, De Villiers JNP, et al. Intronic mutations at splice junctions in the low-density lipoprotein receptor gene. Mol Cell Probes 1999; 13: 257–260.

Pimstone SN, Gagné E, Gagné C, et al. Mutations in the gene for lipoprotein lipase: a cause for hypoalphalipoproteinemia in individuals heterozygous for familial hypercholesterolemia. Arterioscler Thromb Vasc Biol 1995; 15: 1704–1712.

Pohl EH, Kotze MJ, Grant KA, et al. Impact of MammaPrint on clinical decision-making in South African patients with early stage breast cancer. Breast J 2016; 22(4): 442–446.

Press MF, Sauter G, Buyse M. HER2 gene amplification testing by fluorescent in situ hybridization (FISH): comparison of the ASCO-College of American Pathologists guidelines with FISH scores used for enrollment in breast cancer international research group clinical trials. J Clin Oncol 2016; 34(29): 3518–3528.

Raal FJ, Pappu AS, Illingworth DR, et al. Inhibition of cholesterol synthesis by atorvastatin in homozygous familial hypercholesterolemia. Atherosclerosis 2000; 150: 421–428.

Rubinsztein DC, Coetzee GA, Van der Westhuyzen DR, et al. Familial defective apolipoprotein B-100 is rare in South African Afrikaners, Coloureds and Indians. S Afr Med J 1995; 85: 355–357.

Sawe RT, Laing N, Mining SK, Kotze MJ. Experience with the informed consent process and family pedigree assessment for *BRCA1/2* gene screening in Kenyan breast cancer patients. Congress proceedings of the Breast Interest Group of South Africa. Accepted for oral presentation, 21 October 2017a.

Sawe RT, Mining SK, Ofulla AV, et al. Tumor infiltrating leukocyte density is independent of tumor grade and molecular subtype in aggressive breast cancer of Western Kenya. Trop Med Health 2017b; 45(19): 1-11.

Scholtz CL, Peeters AV, Hoogendijk CF, et al. Mutation -59ct in repeat 2 of the LDL receptor promoter: reduction in transcriptional activity and possible allelic interaction in a South African family with familial hypercholesterolaemia. Hum Mol Genet 1999; 8: 2025–2030.

Seftel HC, Baker SG, Sandler MP, et al. A host of hypercholesterolaemic homozygotes in South Africa. Br Med J 1980; 281(6241): 633-636.

Sniekers S, Stringer S, Watanabe K, et al. Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. Nat Genet 2017; 49(7): 1107–1112.

Talmud PJ, Shah S, Whittall R et al. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic hypercholesterolaemia: a case–control study. Lancet 2013; 381(9874): 1293–1301.

Thiart R, Scholtz CL, Vergotine J, et al. Predominance of a 6-bp deletion in exon 2 of the LDL receptor gene in Africans with familial hypercholesterolaemia. J Med Genet 2000; 37: 514–519.

Totzeck M, Mincu RI, Rassaf T. Cardiovascular adverse events in patients with cancer treated with bevacizumab: A meta-analysis of more than 20 000 patients. J Am Heart Assoc 2017; 6: e006278.

Traglia M, Girelli D, Biino G, et al. Association of HFE and TMPRSS6 genetic variants with iron and erythrocyte parameters is only in part dependent on serum hepcidin concentrations. J Med Genet 2011; 48(9):629-634.

Utermann G, Langenbeck U, Beisiegel U, Weber W. Genetics of the apolipoprotein E system in man. Am J Hum Genet 1980; 32(3): 339–347.

Van Cauwenberghe C, Van Broeckhoven C, Sleegers K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. Genet Med 2016; 18(5): 421–430.

Van der Merwe N, Peeters AV, Pienaar FM, et al. Exome sequencing in a family with luminal-type breast cancer underpinned by variation in the methylation pathway. Int J Mol Sci 2017; 18(467): 1-17.

Van Rensburg SJ, Potocnik FCV, De Villiers JNP, et al. Earlier age of onset in Alzheimer's disease patients with both the transferrin C2 and apolipoprotein E4 alleles. Ann N Y Acad Sciences 2000; 903: 200–203.

Van Rensburg SJ, Kotze MJ, van Toorn R. The conundrum of iron in multiple sclerosis - time for an individualised approach. Metab Brain Dis 2012; 27: 239-253.

Van Rensburg SJ, van Toorn R, Schoeman JF, et al. Iron deficiency in two children diagnosed with multiple sclerosis: report on whole exome sequencing. S Afr J Psych 2015; 21(3): 111-112.

Van Toorn R, Schoeman JF, Solomons R, et al. Iron status in children with recurrent episodes of tumefactive cerebral demyelination. J Child Neurol 2010; 25(11): 1401–1407.

Varret M, Rabés JP, Thiart R, et al. LDLR database (second edition): new additions to the database and the software, and results of the first molecular analysis. Nucl Acis Res 1998; 26: 250–255.

Varret M, Rabès J-P, Jore BS, et al. A third major locus for autosomal dominant hypercholesterolemia maps at 1p34.1-p32. Am J Hum Genet 1999; 64: 1378–1387.

Vergotine J, Thiart R, Langenhoven E, et al. Prenatal diagnosis of familial hypercholesterolemia: importance of DNA analysis in the high-risk South African population. Genet Counsel 2001a; 12: 121–127.

Vergotine J, Thiart R, Scholtz CL, Kotze MJ. Clinical versus molecular diagnosis of heterozygous familial hypercholesterolaemia in the South African population. S Afr Med J 2001b; 91: 1053–1059.

Villeger L, Abifadel M, Allard D, et al. The UMD-LDLR database: additions to the software and 490 new entries to the database. Hum Mutat 2002; 20: 81–87.

Villeneuve S, Brisson D, Marchant NL, Gaudet D. The potential applications of Apolipoprotein E in personalized medicine. Front Aging Neurosci 2014; 6: 1-11.

Wall DP, Tonellato PJ. The future of genomics in pathology. Med Reports 2012; 4(14): 1-5.

Whitworth P, Stork-Sloots L, de Snoo FA, et al. Chemosensitivity predicted by BluePrint 80-gene functional subtype and MammaPrint in the Prospective Neoadjuvant Breast Registry Symphony Trial (NBRST). Ann Surg Oncol 2014; 21(10):3261-3267.

Yang C, Chu J, Warren RL, Birol I. NanoSim: nanopore sequence read simulator based on statistical characterization. Gigascience 2017, 6: 1-7.