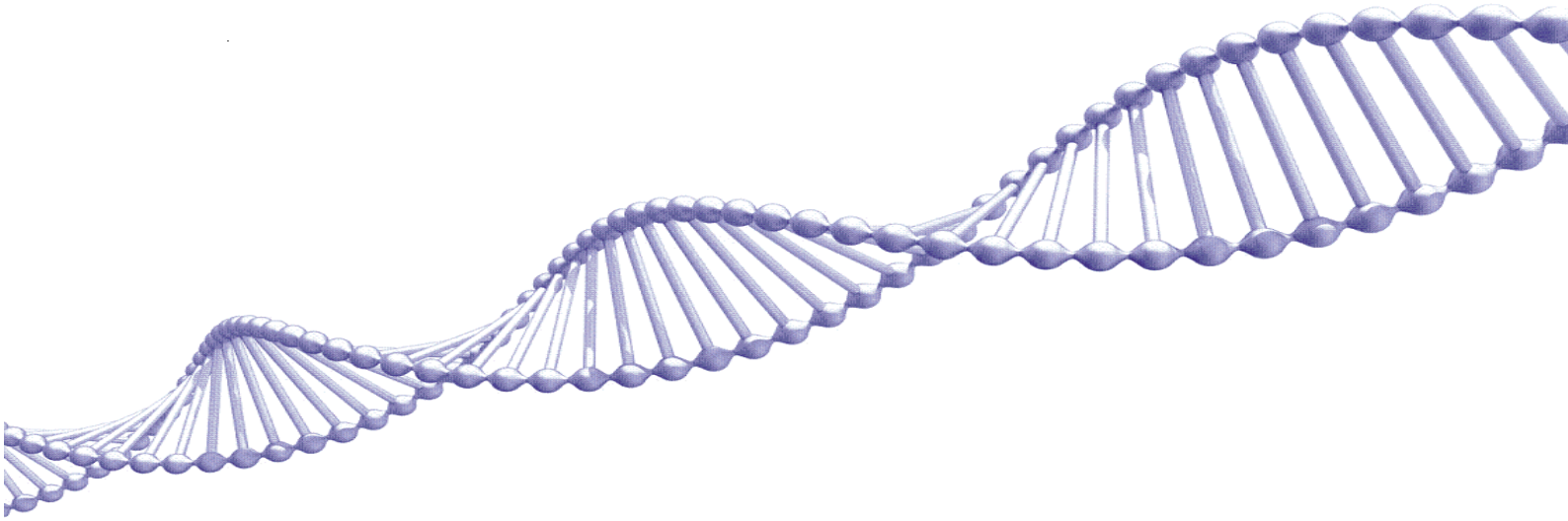
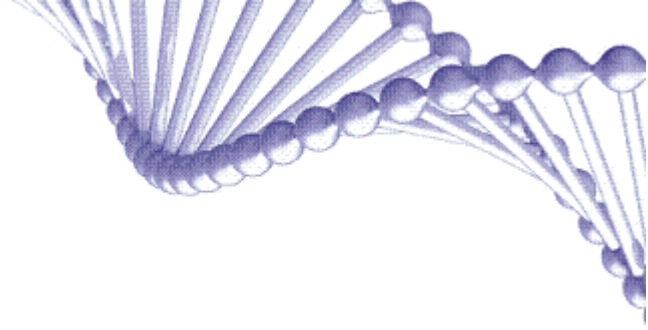


Personalized Medicine

Facts, thoughts and future perspectives

Master thesis

C.A. Meddens



Master thesis
April 2011

Author:	C.A. Meddens, BSc.
Master:	Cancer Genomics and Developmental Biology, Utrecht University
Supervisor:	Prof. Dr. W. de Laat, Hubrecht Institute
Title of thesis:	Personalized Medicine, facts thought and future perspectives

Contents

1. Introduction

Ten years ago President Bill Clinton welcomed Craig Venter and Francis Collins to announce to the world that a new era arrived; the era of the human genome. Both the company 'Celera Genomics' of Venter and Collins' publicly funded 'Human Genome Project' presented their first survey of the human genome. A combination of 3 billion A's T's G's and C's that should be the answer to a great variety of biomedical questions. President Clinton postulated some expectations on new developments in medicine: '*This project promises to lead to a new era of molecular medicine, an era that will bring new ways to prevent, diagnose, treat, and cure disease'*[1].

President Clinton was not the only person at that time with high expectations on new developments in medicine. Just after the first human genomes were published, many scientist wrote articles on the post-genome era and its implications for drug discovery. Patient care was expected to be *revolutionized* through the use of novel molecular predisposition, screening, diagnostic, prognostic, pharmacogenomic and monitoring markers. Diseases would be classified based on their molecular background instead of the clinical symptoms.[2] Furthermore, it was expected that the process of drug development would change from a linear process, in which drugs that failed in either phase of testing were discarded, to a process with many feedback loops. In this new process, information that would emerge in any stage of drug discovery, should have implications both up and downstream phases of the testing process (figure 1).[2] Francis Collins shared the predictions he made in 2000 for the year 2010 with the rest of the world in Nature, one decade after the publication of the human genome. These predictions were: *1. Predictive genetic tests will be available for a dozen conditions. 2. Interventions to reduce risk will be available for several of these. 3. Many primary-care providers will begin to practice genetic medicine. 4. Pre-implantation genetic diagnosis will be widely available, and its limits will be fiercely debated. 5. A ban on genetic discrimination will be in place in the United States. 6. Access to genetic medicine will remain inequitable, especially in the developing world.*[3] Although many people saw lots of opportunities for the post-genome era, the publication of the genomes revealed some complications as well. The 2 parental genomes of C. Venter differed from each other by 0,5%, which was much more than the predicted 0.1% based on single nucleotide polymorphisms (SNPs).[4] This meant that the inter-personal variety was more complex than initially thought. Whereas the variation was found to be larger than expected, the predicted number of genes was much lower than expected. This meant that the complexity could not only be found at the genomic level, but at the epigenetic, transcriptional, translational and post-translational levels as well; levels that cannot be investigated by interpretation of the genomic sequence alone.[5]

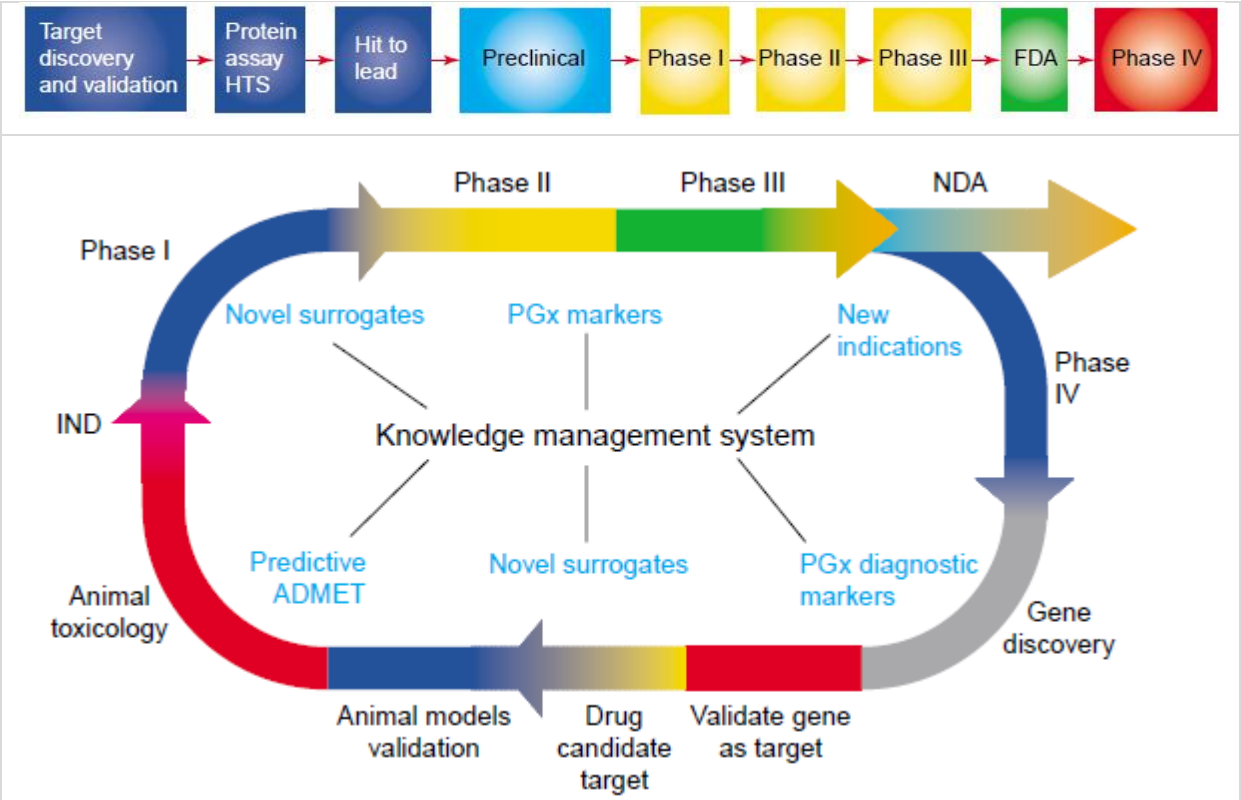


Figure 1. Linear vs feedback process of drug development. *´Genomic information and markers emerging at each stage of the discovery process will be used as tools both upstream and downstream, resulting in better pharmaceuticals and personalized medicine products. A knowledge warehouse will store information enabling continued process and product improvements. Abbreviations: ADMET, absorption, distribution, metabolism, excretion and toxicity; ND, investigational new drug; NDA, new drug application; PGx, pharmacogenomics.'*[2]

Ten years after the warm welcome for Venter and Collins at the White House, we can look back on the first post-genome decade. Which things have changed in- and outside the clinic? Which expectations have become reality and what are the expectations for the next decade(s)? Whereas Collins and Venter needed ten years to sequence one human genome, nowadays a complete genome can be sequenced in one day. As the speed of sequencing is still increasing and the costs are decreasing, the time in which everyone can have their own genome sequenced, lies ahead of us.[6] These 'personal genomes', have led to the development of a new field of research, called 'personalized medicine'. In this field, the goal is to link human genetic variation to physiology and disease. This has great implications for the development of new drugs and the treatment of diseases. Therefore the publication of the human genome was a lead of departure for a new era in which many things have changed, but at the moment most processes are still ongoing. In this thesis several articles on the recent developments in personalized medicine will be discussed. Since these developments are moving 'from bench to bed', the time has come to move one step further and discuss the hurdles that have to be taken in order to incorporate personalized medicine into the clinic.

There are three different ways in which personalized medicine will likely be used by clinicians and patients; First, the prediction of the way a patient responds to a certain drug. Therefore patients might no longer be treated in a trial-and-error manner, but in an evidence based way by using personal genomes as a tool for outcome predictions. Second, the prediction of the risk that a person has to develop disease, possibly leading to prevention. Finally, the determination of the carrier status of parents and therefore the chance of offspring to develop disease. This will influence prenatal (and possibly pre-pregnancy) screening. All these three topics will be discussed in this thesis. I will start with a few examples of how personalized medicine is already being used inside and outside the clinic. Then, current research on drug response will be discussed. Thereafter the research on predictions on disease risk and carrier status are presented. In the final chapter several issues will be discussed, both practical and ethical, that researchers and clinicians will come across when personalized medicine is implemented into clinical practice.

2. Current personalized medicine

Although the development of personalized medicine has only just begun, there are already several examples of the way treatments are adjusted based on predictive factors. In this chapter the treatment of breast cancer will be discussed, in order to show an example of how treatments were personalized in the pre-genomic era. By using breast cancer as an example, the way treatments in the clinic have changed in time, first based on macroscopic and later on microscopic factors, can nicely be illustrated. Something that is done in order to diagnose patients in a pre-clinical stage, is neonatal screening. Therefore neonatal screening is used here as an example of a pre-genomic-era risk prediction. As the use of disease risk predictions based on genetic information is still limited inside the clinic, the first post-genome risk predictions are currently found outside the clinic. There are several companies that offer predictions on a customers' personal risk to develop diseases or the way they will respond to certain therapies, based on their DNA sequence. These direct-to-customer risk predictions will be discussed in the last section of this chapter.

2.1 Personalized medicine in the pre-genomic era

Breast cancer

Breast cancer is the type of cancer with the highest incidence in females in USA in 2010 (28% of new female patients with cancer suffer from breast cancer)[7]. From the moment a malignant tumor is diagnosed until the moment the patient is either cured or undergoing palliative treatment, many clinical decisions have to be made. These decisions concern the type, combination and the severity of treatment. Already since the early days, the macroscopic aspects that were used to determine the type of treatment were histology of the tumor, the tumor size, the axillary lymph node status and whether the tumor had metastasized to a distant location. Lumpectomy (surgically removing the tumor) is performed in patients with a small tumor (diameter < 5 cm), whereas mastectomy (surgically removing the whole breast) is performed in patients with a large tumor (diameter > 5 cm).[8] During surgery lymph nodes are being removed, after which tissue is analyzed for the presence of lymph node metastasis. This information is then used to decide whether or not to start with adjuvant chemotherapy (solely patients with a positive lymph node status would be advised to start chemotherapy). Although tumor size and lymph node status have a prognostic value for the outcome of treatment and the overall survival, these factors have a low sensitivity and specificity. This means that the number of relapses and the number of patients that unnecessarily undergo mastectomy is

relatively high.[9] Therefore clinicians have continued to search for tumor characteristics that have better prognostic value for the outcome of certain therapies.

This has resulted in a few other factors that are nowadays used to determine adjuvant treatment of breast cancer. After surgery, the tumor tissue is stained for the expression of estrogen and progesterone receptors (ER and PR).[10] Already before the finding that some of the breast tumors express these receptors, patients were treated with drugs that inhibit the function of these receptors by competitively binding their steroid binding domain. Since all malignant breast tumors are analyzed for the expression of ERs and PRs, all patients with tumors that express estrogen or progesterone receptors are treated with 'adjuvant endocrine therapy' (usually tamoxifen) combined with cytostatics, whereas patients with a negative receptor status are treated with cytostatics alone. Because of the development of receptor expression analysis and the adjuvant endocrine treatment the number of relapses and treatment side-effects have decreased. Firstly, by treating patients that are known to respond to such therapy, the disease free survival and the overall survival of these patients have increased. Secondly, since only patients that express the receptors are treated, the other patients are no longer exposed to treatment side effects.[11] Recently another predictive factor has been identified; the overexpression of the human epidermal growth factor receptor Her-2/neu by the tumor tissue is associated with aggressive tumors and poor prognosis. This has lead to the development of a new drug for patients of which the breast tumor tissue overexpresses Her-2/neu. It is a monoclonal antibody against the extracellular domain of the Her-2/neu receptor (trastuzumab). Treatment with trastuzumab increases the disease free survival and the overall survival of this specific group of patients.[12]

We can conclude that currently, many prognostic and predictive factors are used in the treatment of breast cancer. Not only tumor size and (axillary lymph) metastasis are taken into account, but also the expression status of estrogen and progesterone receptors and the overexpression of Her-2/neu are used to determine the right treatment for the individual patient. But can we speak of personalized medicine already? The number of groups to which the patients are classified is still very small and number of patient per group is very large. Furthermore, non-responders can still be found in each group. Therefore the treatment of breast cancer has to be further personalized, in order to treat patients in an evidence based manner. Since the human genome was published, many genes have been annotated and it became possible to create expression profiles in which the expression levels of many different mRNAs could be measured simultaneously. This method was used to study the expression profiles of breast tumor tissue, after which statistical methods have been used to determine subsets of genes that have a predictive value concerning the treatment and prognosis of

breast cancer[13]. In the next chapter, this post-genome high throughput analysis and its implementation in the clinic will be further discussed.

Neonatal Screening

Risk prediction and early diagnosis of disease has already been a goal in medicine for a long time. In the post-genome era, people are trying to achieve this by linking genetic variants to a disease phenotype. But already long before the human genome was sequenced, early diagnostic methods were (and are still being) used. There are several diseases that can be diagnosed in newborns. For a selection of these diseases, early start of treatment results in a better outcome for the patients. Therefore in many countries newborns are screened for a limited number of diseases, that are all selected because early treatment for these diseases is available.[14] Although the treatment for these diseases is not 'personal', the approach in which blood is screened for biological markers and patients are diagnosed before the first symptoms are present, is something that is likely to become the standard for many diseases once everyone's genome can be sequenced. Currently, many countries in Europe and the USA have their one screening program, each with a distinctive set of diseases that are tested in newborns. Almost all countries screen for phenylketonuria and congenital hypothyroidism, two diseases that can be easily treated. When these diseases are not treated, both diseases cause mental retardation, something that can now be prevented.[15] In order to show how screening is performed and incorporated into a national screening program, cystic fibrosis screening is used as an example here.

Recently, a new test in which Cystic Fibrosis (CF) can be diagnosed, was added to the screening program in the Netherlands. CF is an autosomal recessive disease, in which the two CFTR genes, encoding the cystic fibrosis transmembrane conductance regulator, are mutated. This results in a non-functional (or absent) chloride and bicarbonate transmembrane transporter. As a consequence, the electrolyte content of the fluid that covers several epithelial membranes, is changed. In the lungs of CF patients this causes a thick viscous mucus, that impairs the mucus clearance from the upper airways.[16] It has been shown that early treatment of CF results in less pulmonary damage and improved growth and survival.[17] Therefore all children in the Netherlands are now screened for CF. After birth a blood sample is taken from the heel of the newborn. First, the concentration of immunoreactive trypsinogen (IRT) in measured. This is elevated in CF patients due to a decreased pancreatic function. Therefore, elevated IRT is the first result that points towards CF. In all positive samples, the two CFTR genes are sequenced and screened for a panel of known CFTR-mutations.[18] In this way, the knowledge on the pathogenesis of this 'single gene'-disease, is used to screen for CF patients and start treatment in newborns.

We have seen two examples of how current knowledge on disease is used to improve diagnosis and treatment of disease. In the next section, we will show how personalized medicine is currently used outside the clinic, in a commercial manner.

2.2 Commercial direct-to-costumer SNP analysis

As we have seen in the previous sections of this chapter, it is possible to screen an individual for characteristics of disease or carrier status. To search for (disease) characteristics based on someone's genotype, all you need is a person's DNA and the tools to analyze it. In the last few years, new companies have arisen that provide people information on their own genome. One of this companies is 23andme.

When someone sends 23andme a sample of saliva, this person's genome is screened for inter-human variants, called single nucleotide polymorphism (SNPs). SNPs are DNA variants of a single nucleotide that commonly differ between individuals. SNPs can be found all over the human genome and are mapped by the International Hapmap consortium, in the Hapmap project.[19] This map of SNPs is used in many studies to link these variants to diseases, carrier status, human traits and drug response.[20] As these studies continue to be published, individuals can constantly be updated on their variant status. At 23andme, 1.000.000 SNPs are analyzed, by using a chip that contains probes for each SNP. This information is then used to inform the customer about their own traits, disease risk and drug response based on what is known from literature. Each month newly published findings are linked to each genome and the customer is

Figure 2.1 The workflow that is used by 23andme is depicted here.

updated on what has been found (figure 2).

For $199, 23andme performs a SNP-analysis and keeps their customers informed on four subclasses of information; disease risk, carrier status, drug response and ancestry. To get an overview of the information such companies can offer their customers, an example of each subclass is discussed here. At this moment the company includes information on 95 diseases. As an example, the information related to bladder cancer that is provided is based on four separate studies in which three SNPs (rs798766, rs9642880, rs710521) have been linked to increased or decreased chance to develop bladder cancer.[21-24] Therefore the variants of the customers genome is used to give information about the relative chance the individual has to develop bladder cancer, based on these three SNPs.

There are 24 diseases on which the customer is informed about whether or not he or she is carrier for this disease. A carrier is someone that has one of the two disease alleles, but does not have the disease phenotype. Bloom's syndrome is one of these diseases, caused by mutations in the BML gene. Patients that are affected with this autosomal recessive syndrome are exceptionally small throughout life and suffer from many complications, especially cancer. The predisposition to develop cancer is due to the defective BML recQ helicase homolog protein that results in chromosome breakage.[25] Customers that are found to be carrier of Bloom's syndrome could use this information to prevent getting children with Bloom's syndrome, as their partner could be screened for mutations in the BML gene as well.

Customers are also provided with information on how they likely respond to 19 drugs. An example is Metformin, a drug that is used in treatment of type II diabetes. It is the most prescribed drug in patients with type II diabetes, but the effectiveness of Metformin varies a lot between individuals. As under- or over-treatment with Metformin can give severe side effects (hyperglycemia and hypoglycemia respectively), information on variants that influence response to Metformin[26] could be used by clinicians to adjust initial Metformin dose in individual patients.

As modern customers do not only want to be provided with all this serious information, 23andme provides more trivial ins and outs based on an individual's genomic data as well. For example, customers can discover the global origins of their ancestors and compare their genome to genomes of 50 populations worldwide. Furthermore, it is possible to find out what your odds are for male pattern baldness or what type of earwax you have (wet or dry), of course again based on a scientific article in Nature Genetics.[27]

In this chapter, I have given an (incomplete) overview of the current state of personalized medicine in the clinic. It is starting to be used to predict a patients drug response (currently, mostly in cancer

treatment) and it is used to early detect diseases in newborns. Furthermore I have illustrated what kind of information can be provided to customers by commercial companies. In the final chapter of this thesis I will discuss the possible implications of ongoing commercial sequencing for clinicians. In the next two chapters new developments in personalized medicine will be discussed and by sketching several examples, I will give an idea of what the near future might look like.

3. Drug response

Since the publication of the human genome, research concerning the development of new drugs has changed tremendously from the development of 'one size fits all' drugs to medication that can be prescribed based on a patients individual genetic background. Although the incorporation into the clinic is lacking behind, these developments are quickly progressing. This is due to the possibility to retrieve information on a patients' genome sequence, that can be used to determine individual characteristics involved in drug response. This is called 'genomic medicine' or 'pharmacogenomics'. Genomic medicine is defined as the use of information from genomes and their derivatives (RNA, proteins and metabolites) to guide medical decision making[28], both improving patient outcome and minimizing adverse effects[29]. In this chapter a few examples of pharmacogenomic research will be discussed. I will first show how prescription of drugs can be based on genetic variants and second I will show how tumor characteristics can be used to predict prognosis and metastasis, which can be used in therapy decision guiding. Thereafter, the discussion will focus on the challenges that have to be faced in order to implement medical genomics into the clinic.

3.1 Pharmacogenomics and ADME genes

Currently, many drugs are prescribed to large groups of patients, while the desired effect is only achieved in a limited number of that patients. This means that many patients do not respond to the prescribed drug. Furthermore, a number of drugs cause side effects in part of the patients that use them. These side effects can be so severe that some patients cannot continue using these drugs. For most drugs the molecular mechanism of action is unknown and it is not possible to predict if a certain drug will be effective in a patient or if the drug is likely to cause severe side effects. However, there are some factors that are common in the kinetics of all drugs and these factors might be used in the future to predict drug response in individual patients. Although these common factors are not the only factors with predictive value for drug response, they are discussed here because many promising studies have been done on this specific set of genes. [30]

All drugs, regardless of their function, first have to be absorbed, after which they are distributed, metabolized and finally from the body or from the side of action. In most pharmacokinetic reactions these four processes, Absorption, Distribution, Metabolism and Elimination (together called ADME), are guided and/or performed by a limited number of proteins. Inter-individual variation in proteins and genes that are commonly involved in those processes can explain part of the variation in drug response. For example, a mutation in a protein that absorbs a certain medicine from the ileum could cause variation in the efficiency of this process and will therefore result in variation in concentrations

of the drug in different patients. ADME genes are usually grouped into four categories; 1. Phase I and II metabolizing enzymes, that are respectively responsible for modification of functional groups and conjugation of endogenous moieties. 2. Transporter proteins, which are responsible for transport in and out of cells. 3. Serum binding proteins, which bind molecules in order to transport them trough the circulation. 4. Transcription factors and modifiers that are involved in the transcription of other ADME genes.[30] Genetic variation in many ADME genes have been intensively studied.[31,32] Therefore it is has become possible to link this genetic variation to variation in drug response and use the outcome of such studies as a guidance for the prescription of drugs.

When it comes to the prescription of drugs, it is important to notice that scientific knowledge on genetic variants that cause variation in drug response does not automatically lead to implementation of this knowledge into the clinic. The use of such information for medical decision making has to be approved by national institutions. In 2005 the US Food and Drug Administration (FDA) has made it possible to submit pharmacogenomic data together with new pharmaceutics, as the FDA anticipated on an increase in genomic data in drug development and response.[33] Therefore it is no longer a scientific exercise to couple drug response to genetic variants, but it is becoming practice to administer drugs with information in the drug label on genetic variants and drug response. In the future it is likely that determination of differences in drug response caused by genetic variation, will become incorporated into the clinical testing phases. Currently, the most successful example of a drug for which genetic variant in ADME genes are causally related to drug response is Warfarin.

Warfarin

Warfarin is the most prescribed anticoagulant worldwide. Anticoagulants are used in patients that are at risk of developing infarctions caused by thrombi or emboli. By using these drugs, a part of the process of coagulation is inhibited, making the process as a whole less efficient. As Wafarin inhibits coagulation, an undesired side effect that is that patients have a higher change of developing hemorrhages. In individual patients the range of doses at which Warfarin is effective, but not toxic (the dose at which it doesn't cause bleeding) is very small. This means that it is important that patients get the exact right dose, in order to prevent infarctions or hemorrhages. As there is an 20-fold inter-individual variation in required Warfarin dose, predictive factors of the required dose are needed.[34] Until recently, the only variables that could be used to predict the required dose were clinical and environmental factors like age, foods, interactions with other drugs and gender.[35] But nowadays genetic variants have been discovered that are predictive as well.

Warfarin is a racemic mixture of R- and S-Warfarin, with S-Warfarin having a 3 to 5 times more potent anticoagulant effect[36]. Variants in enzymes and proteins that are involved in the pharmacokinetics and pharmacodynamics of S-Warfarin are therefore likely to contribute to the interpatient variation in Warfarin response. S-Warfarin is almost exclusively metabolized by the enzyme CYP2C9 to 7-hydroxywarfarin in the liver. 11 Genetic polymorphisms have been described for the CYP2C9 gene, of which CYP2C9*2 and CYP2C9*3 are the most frequently occurring variants. In various studies in different populations these variants were indeed found to account for part of the variable dose requirement of Warfarin. Patient that are hetero- or homozygous for CYP2C9*2 or CYP2C9*3 need lower Warfarin dose and were found to have an increased risk of bleeding complications[37].

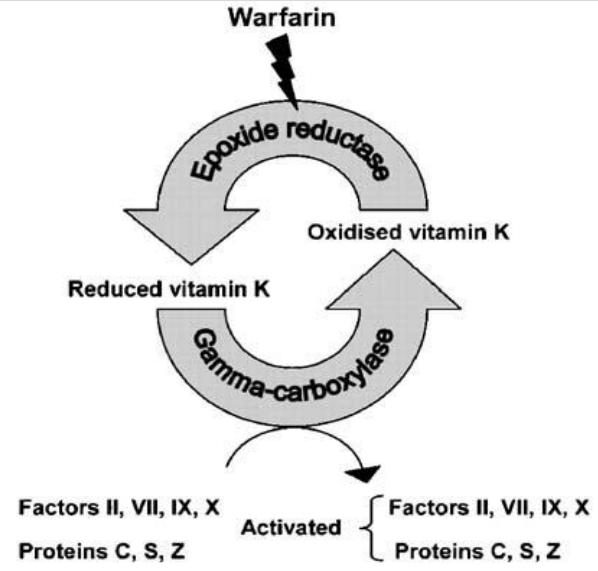


Fig 3.1 Mechanism of action of Warfarin.[38]

As the target of Warfarin is known, it is possible to search for interindividual variation in the pharmacodynamics of Warfarin, rather than only in the pharmacokinetic ADME-genes. Warfarin performs its function by inhibiting Vitamin K epoxy reductase complex subunit 1 (VKORC1) and therefore limits the regeneration of reduced Vitamin K. Reduced Vitamin K is a cofactor for the activation of several clotting factors, thus Warfarin indirectly inhibits the activity of these clotting factors (and therefore secondary coagulation).[38] Genetic variants could cause variance in the effectiveness of the inhibition of VKORC1 by Warfarin. It has been shown that in a group of 201 patients that use Warfarin, the dose response (measured by INR (international normalized ratio for prothrombin time)) correlates with common SNPs in the VKORC1 gene, meaning that the required dose could be predicted based on the variant. As these SNPs were found in non-coding regions, they are thought to affect the transcription, splicing or stability of VKORC1.[38] In the first Genome Wide Association Study (GWAS) that was performed to detect polymorphisms that alter drug response, Takeuchi *et al.* have validated that CYP2C9 and VKORC1 are associated with required dose of Warfarin. They found that CYP2C9 and VKORC1 respectively account for 30% and 12% of the Warfarin dose variance. Furthermore a SNP at CYP4F2 was found to account for another 1,5% of the dose variance.[34] As genetic tests for these SNPs are widely available, it will soon be possible to retrieve information about the SNPs of an individual

patient and use this information together with clinical and environmental factors to determine the required dose for each patient.

3.2 Personalized medicine in cancer

Cancer research is currently the leading field in personalized medicine. There are two major differences between personalized medicine in cancer and in other diseases though. First, the factors that are used to distinguish different groups of patients are tumor-specific factors, rather than somatic genetic variation that is used in other disease. Second, personalized medicine approaches in cancer are currently mostly used to determine the prognosis of the disease, while in other fields people are looking for factors that can tell whether a drug will be able to perform its function in a certain patient. This is due to the fact that the major factor that is used to guide decision making in cancer therapy is the prognosis of the course of the disease. An example of a method in which variation in mRNA expression is used as an prognostic factor is the Mammaprint. In the last section of this chapter a non-genomic, but very innovative, predictive factor will be discussed.

Mammaprint

In the previous chapter we have seen how tumor specific factors are currently used to determine whether adjuvant systemic therapy is needed in treatment of breast cancer and what types of therapy can be used. At this moment, the major reason to start chemotherapy after surgically removing the tumor is the presence of distant metastases or positive lymph nodes. During surgery, the lymph nodes that are directly in contact with the tumor are removed. Therefore it is possible to search for microscopic disease in the tissue of these lymph nodes. Determining the presence of distant metastases is not as easy, as the tissue to which the tumor has metastasized is not known. Furthermore, a tumor (or metastasis) is currently only clinically detectable when its volume is minimally 1 cm^3, therefore it can only be found when it consists of at least 10^9 cells. Chemotherapy reduces the risk of distant metastases by 33%, but it is estimated that 70-80% of the patients would survive without chemotherapy. Ideally, the patients that would benefit from the treatment should be classified on forehand, especially since chemotherapy has a high morbidity and can even cause death.[39]

In the near future it will be possible to predict whether a tumor has the potency to metastasize, based on tumor specific characteristics. Van ‘t Veer *et al.* have developed an mRNA expression signature that is predictive for a short interval (< five years) to distant metastasis.[40] This signature was developed by using an hierarchical clustering algorithm to cluster the microarray expression profiles (5000 genes) of 98 primary breast cancers. It was found that the tumors could be clustered in

two groups. As the study was performed in retrospect, the clinical outcome (development of metastasis within five years) was known for these patients and it was possible to determine which genes were predictive for the development of metastases. This group of genes consists of 70 genes and is called the Mammaprint. In another study this expression signature was tested on 295 patients and was found to be very predictive for the clinical outcome of these patients as well.[41] Furthermore, statistical analysis showed that the signature has a higher specificity than clinicopathologic factors and that it can be used as an independent factor in prognostic risk assessment.[42] By using the Mammaprint, patients that are at risk for the development of metastases can get aggressive chemotherapy, whereas patients that are not at risk will no longer be treated with cytostatics.

Circulating tumor cells

Expression profiles of tumor tissue is possibly not the only method to predict metastases in cancer patients. During the process of metastasis, tumor cells detach from the primary tumor and move either through blood or the lymphatic system, to a distant organ.[43] Quantification of the amount of Circulating Tumor Cells (CTCs) could therefore be a measure for the amount of cells that are detached from the tumor and indirectly for the process of metastasis. Although the clinical relevance of CTCs is still subject of debate, the detection of CTCs as an method to determine prognosis is discussed here, as it shows the diversity in development of methods in the field of personalized cancer care.

The dissemination and the implantation of tumor cells at distant sites is a multi-step process. After dissemination the tumor cell circulates through the body, during this part of the process most cells die, only cells that perform well can succeed in invasion of a distant organ. Figure 3.2 depicts a range of processes that lead to the development of disseminated malignancy and the obstacles that disseminated cells have to overcome in order to successfully occupy a distant organ.[43] There is evidence that CTC detection can be used for prediction of cancer prognosis and the diagnosis of metastases in breast cancer patients. It has been shown that the detection of CTCs in patients with a negative lymph node status is an independent prognostic factor for early clinical relapse and disease-related death.[44] In another study, the quantification of CTCs was used to evaluate the response of breast cancer patients to systemic chemotherapy. A clear correlation was found between a decrease in number of CTCs (more than 10-fold) after systemic therapy was started and the final response of the tumor. Therefore, CTC detection might be a tool for closely monitoring therapy response.[45]

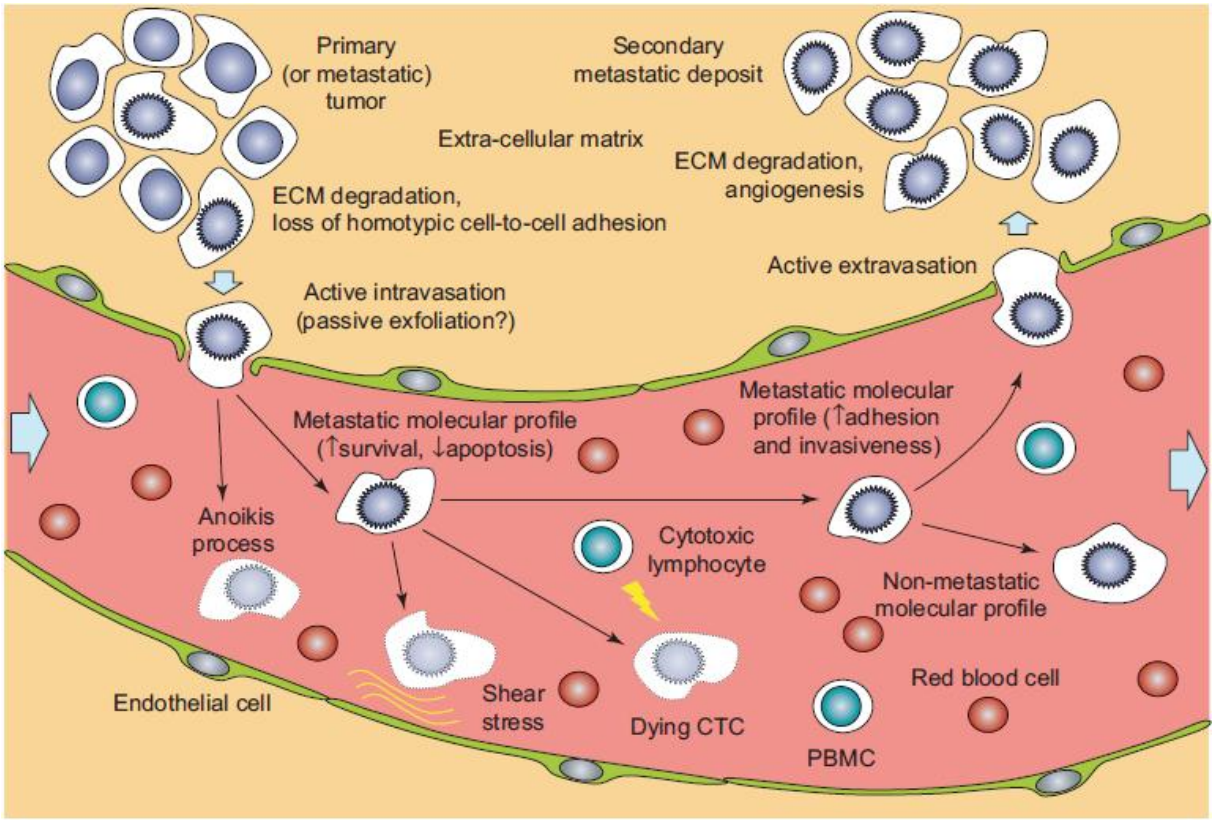


Figure 3.2 The seed and soil hypothesis. In this figure all steps that tumor cells have to go through in order to invade a distant target, are depicted.[46]

Before CTC detection can be incorporated in decision guidance for cancer therapy, the technical procedures will have to be improved. The detection of CTCs can be done in two different ways, first by Polymerase Chain Reaction (PCR), in which primers for tumor specific DNA or mRNA targets is used. And secondly by cytometric techniques, in which cell morphology and cell surface markers can be used to detect CTCs.[46] The limitations of the detection are the limited number of known tumor specific markers that can be used, complicated by the fact that these markers should only be present in tumor cells (not on other circulating cells) and ideally are shared between CTCs of different patients with the same type of cancer. Furthermore, the enrichment methods for CTCs are still limiting in detection of CTCs with high sensitivity and specificity. Several methods for CTC detection have been described so far. Peripheral blood can be filtered, in order to separate the small leukocytes from the larger tumor cells. Furthermore, CTCs can be enriched by tumor specific antibodies that can be bound to a magnetic microbead after which tumor cells can be magnetically selected. One of the problems with all enrichment methods is that it is hard to control for inter-sample differences.[46]

We have seen that in personalized cancer care, methods are being developed that help to predict the prognosis of the disease and especially the metastatic status. In the previous section, it was shown that genetic variation in ADME genes can be used to predict the way an individual responds to a certain drug. It should be noted that the use of genetic variants in ADME genes is used for the response to anti-cancer agents as well. For example, the efficiency of metabolism of tamoxifen was found to be related with genetic variation in the CYP2D6 gene and a mutation in the DPD gene (DPD*2A) causes the inability to degrade fluorouracil (5-FU) and is therefore associated with lethal outcome of treatment with 5-FU.[47] We can conclude that many new methods to predict drug response and guide medical decision making are being developed and will be used by clinicians in the near future.

4. Health risk assessment

For most diseases it is not possible, based on predictive factors, to be 100% sure if someone will, or will not develop a certain disease. Health risk assessment can then be used to determine what risk an individual person has to develop disease. This information can be used to take preventive measures, start early treatment or even to support a diagnosis. Risk factors can be both genetic and non-genetic. Non-genetic factors, like environmental factors, are already studied and used for risk assessment for many decades. As genetic factors, which are being discovered since the post-genome era, will be a replenishment to non-genetic factors, both will be discussed in this chapter.

4.1 Health risk assessment in current practice

Knowledge on the risk an individual patient has to develop a certain disease can help clinicians to prevent and diagnose disease. Currently, different types of risk factors are used, especially by general practitioners, as support for diagnosis and to guide medical decisions.[48] Risk factors can be divided into different groups: family history, personal factors, biomarkers and environmental factors. All of these factors are determined by epidemiological studies, in which the presence of these factors is correlated to the prevalence, incidence or prognosis of disease. As cardio-vascular disease (CVD) is the main cause of death in the developed world, many studies have been performed to determine risk factors. Therefore, I will use CVD as an example to show how different groups of risk factors are currently being used to assess the risk of an individual patient to develop a certain disease.

Cardio-vascular disease

Several studies have determined family history of coronary heart disease as an individual risk factor for developing the disease[49, 50]. These studies did not determine the genetics of CVD, but they investigated whether offspring of people that suffer from CVD have a higher risk to develop CVD. The knowledge that a family history of CVD increases the risk of developing CVD is used by clinicians; a patient with a positive family history is informed about the environmental risk factors, so the patient can prevent the development of CVD by lifestyle changes. Furthermore, general practitioners are more alert in noticing the first symptoms of CVD in these patients.[51] Therefore, this risk factor is used for prevention of, and screening for CVD. When it comes to prevention of CVD, environmental factors that cause CVD come into play. Tobacco use is the environmental factor that contributes most to the development of CVD. In women under the age of 50 years, smoking is the leading cause of coronary heart disease.[52] Furthermore, obesity and physical inactivity can cause atherosclerosis and therefore predispose a person to develop CVD.[53] Risk factors that cannot be prevented are the personal factors. In CVD age and gender are most distinguishing. The risk of developing CVD increase

with age. Furthermore, men and postmenopausal women have a higher risk as well[54]. Whether this is due to the protective effect of estrogen is still under debate[55]. There is no real biomarker for CVD, but there are markers that are associated with a higher risk to develop CVD and there are markers that are used to diagnose a recent myocardial infarction. A marker that is associated with increased risk is a high level of Low Density Lipoprotein (LDL) (relative to HDL). This finding has led to the development of pharmaceuticals that bind the LDL receptor (statins) and therefore lower the risk to develop CVD.[56] The markers that are used for the diagnosis of recent myocardial infarction are creatinine-kinase MB and Troponin-I and –T. These are cardiac specific markers that enter the circulation after necrosis of cardiomyocytes. The quantitative measurement of these factors can be used as an indication of the size of the infarction.[57] In the next sections we will see that besides all the factors that are currently used for health risk assessment, tools are being developed in order to use mutation and variation analysis as a new factor in the post-genome era.

4.2 Health risk assessment in the post-genome era

With the continuing drop in sequencing costs and the $1000 genome coming closer, a time will come in which it is possible to sequence everybody's genome. As we have seen in the previous chapter, genomic variation can be used to predict drug response. Since many diseases are caused by mutations and genomic variants, information on the genome of an individual can be used to predict whether someone will develop a certain disease or has an increased chance to develop a disease. In case of Mendelian diseases there is no real risk assessment involved, as it is possible to determine whether or not an individual has the disease. In contrast, for complex diseases this is not as easy and therefore genetic factors will play a role in risk prediction rather than in diagnostics. Health risk assessment can be useful when early treatment has a positive effect on the progression of disease, or in the case of an increased risk, preventive measures can be taken. Sequencing of a genome is not the only thing that has to be done in order to predict or diagnose disease, as sequence information is useless without knowing which variants and mutations cause the disease or increased risk. Below I will give an overview on the research that is done on both Mendelian and complex diseases and the impact this will have on personalized medicine.

Mendelian diseases

Mendelian or monogenic diseases are diseases that are caused by mutations in one gene. Some of these diseases are only seen in homozygotes (recessive) whereas others can be found in heterozygotes (dominant). Monogenic diseases are rare and are therefore caused by rare variants[58]. In most diseases with a Mendelian inheritance pattern the gene that is responsible for the phenotype is known. These genes are mostly found through linkage studies. In this method the location of the

gene in the genome is determined in families in which the phenotype segregates, by comparison of the inheritance of a set of polymorphism between the affected and the non-affected family members.[59] Once the gene that causes a Mendelian disease is known, this information is used as a tool to diagnose both patients and carriers. Furthermore, it gives new insights into the pathogenesis of disease and can be used for development of new pharmaceuticals.

There is a large spectrum of monogenic diseases, ranging from familial cancers to enzyme deficiencies. One example of a monogenic enzyme deficiency is Gaucher disease, which is a recessive disease in which the β-glucocerebrosidase gene is mutated. It causes lysosomal storage of glucocerebroside, that leads to neurologic dysfunctions, bone malformations and hepatosplenomegaly.[60] Another enzyme deficiency is Cystic Fibrosis, as is discussed in chapter 2. An example of a monogenic familial form of cancer is Familial Adenomatous Polyposis (FAP), that leads to the development of many tumors in the colon. It is caused by a mutation in the APC (adenomatous polyposis colon) gene and although both genes need to be mutated in order to develop the disease, the inheritance pattern is autosomal dominant. This has to do with the fact that the mutation of the unaffected allele almost always occurs in patients with one affected allele.[61]
All information on known Mendelian traits and diseases is well documented and freely available online at the 'online Mendelian inheritance in man' OMIM-database (www.ncbi.nlm.nih.gov/omim).[62] In figure 4.1 the information that is currently provided by the database is shown.

	Autosomal	X-Linked	Y-Linked	Mitochondrial	**Total**
* Gene with known sequence	12605	620	48	35	13308
+ Gene with known sequence and phenotype	314	18	0	2	334
# Phenotype description, molecular basis known	2725	236	4	28	2993
% Mendelian phenotype or locus, molecular basis unknown	1632	134	5	0	1771
Other, mainly phenotypes with suspected Mendelian basis	1831	130	2	0	1963
Total	19107	1138	59	65	**20369**

Figure 4.1 Information available at the OMIM-database as of April 3th 2011.[62]

The determination of genes responsible for the phenotypes of monogenic disease is something that was already done a long time before the first human genome was published. Although many knowledge was already available in the pre-genome era, things will change in the post-genome era. When everybody will have his or her genome sequenced, the screening of newborns will contain many more diseases then the limited number that is screened for nowadays. Patients can be informed before disease onset, even if there is no early or preventive treatment available. The ethical problems that are involved when providing this information to patients are discussed in chapter 5. Many Mendelian diseases are homozygous diseases, meaning that both alleles of the gene need to be mutated in order to develop the disease. When people have their whole genome sequenced not only homozygous mutations will be found, but also the cases in which only one allele is mutated will come to light. These people will not develop the disease but are so called carriers. When two carriers of the same disease would have a child together, the chance is 25% that their child will inherit both mutated alleles and develop the disease. Therefore this knowledge will bring about another ethical problem: when should people be provided with this information and should it be possible to use this information to screen for embryo's that do not have both mutations? These questions will be discussed in chapter 5 as well.

Complex diseases

Complex diseases are common diseases caused by a combination of environmental and genetic factors and they are therefore sometimes called polygenic or multi-factorial diseases. The genetic background of these diseases is complex, as many different genes play a role and it seems that a combination of many variants causes the phenotype. The combinations of variants that cause the disease even differ from patient to patient.[63] Whereas the risk prediction for Mendelian diseases is straight forward (you have either zero, one or two of the disease alleles), risk prediction for complex disease is very complicated due to the spectrum of factors that play a role. In one of the previous sections we have seen that the influence of environmental factors on the development of disease is investigated by epidemiological studies. Since it is possible to study whether patients have been significantly more frequent exposed to a specific environmental factor compared to a group of controls, the same can be done to see if a specific genetic variant is overrepresented in a group of patients. This is what is done in a high-throughput manner, called genome-wide association studies (GWASs). In these studies, over 500,000 common variants are studied in both a group of patients and a group of controls, in order to determine variants that are linked to complex diseases like for example rheumatoid arthritis, cardio-vascular disease and Parkinson's disease.[64, 65] The variants that are used in GWASs are usually single nucleotide polymorphisms that are common (>5%) in the

human population. This has i. a. to do with the hypothesis that common diseases are caused by common variants (common disease, common variant hypothesis), which argues that *'genetic variations with appreciable frequency in the population at large, but relatively low penetrance (or the probability that a carrier of the relevant variant will express the disease), are the major contributors to genetic susceptibility to common disease'*[66]. An example of the outcome of a GWAS is shown in figure 4.2, in which all human chromosomes are depicted.

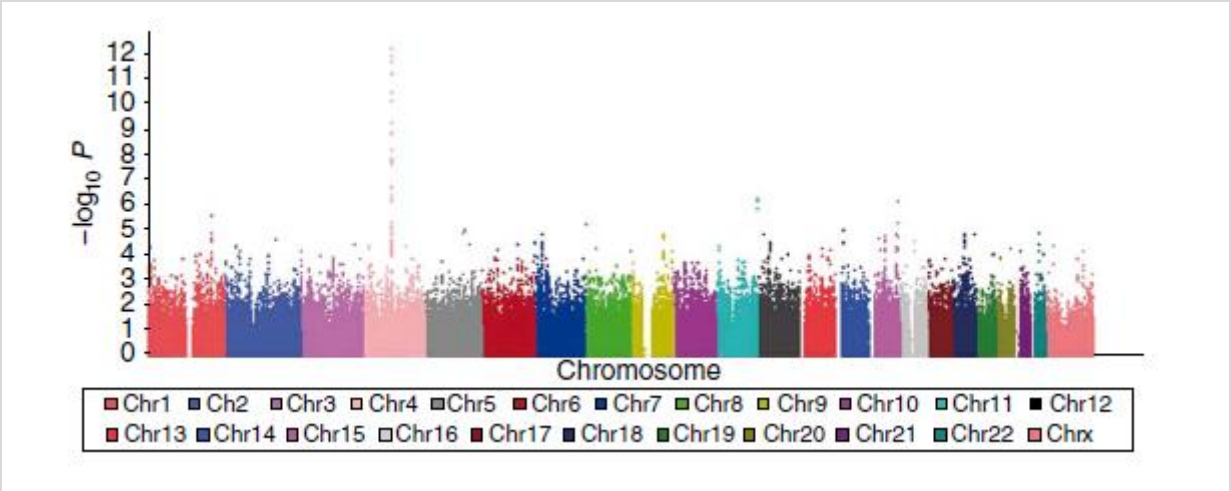


Figure 4.2 Results of a GWAS. On the x-axes all variants that are studied are represented according to the chromosome. On the y-axes the P-value of association to the disease is depicted. On chromosome 4 one variant shows a highly significant association with the disease.[65]

The first GWAS was published in 2005 by Klein *et al.*. In this study, age-related macular degeneration was found to be associated with an intronic SNP in the gene that is coding for Complement factor H. People that are homozygous for the risk allele were found to have an odds ratio of 7.4 (P value $<10^{-7}$, compared to people that are homozygous for the normal allele).[67] Since 2005, 450 genome-wide association studies have been published. The completion of the International HapMap project in 2007 was another step forward, as this is a source of all known human genetic variants.[63] Whereas the first published GWAS was very successful in determining a variant that has great predictive value for a specific disease, most other studies were not able to elucidate such high risk factors. In most studies, the variants that were found to be significantly predictive for the risk to develop disease had odds ratio's between 1.2 and 1.3, meaning that these variants are only responsible for a minor increase in risk.[68] Therefore, both computational and technical improvements are being developed, in order to determine all variants that are associated with complex diseases. Furthermore many institutes are starting to collaborate in large consortia in order to increase the statistical power of the studies, which will lead to new insights on the genetic background on multi-factorial diseases.[69]

It is expected that in the next decade, the number of GWASs that are published will increase exponentially, but will probably be based on whole genome sequencing rather than SNPs only. This will lead to an enormous increase of information on variants that are associated with multi-factorial diseases. Whereas new knowledge on Mendelian genes and diseases can easily be incorporated into the clinic for diagnostics, for polygenic diseases this is more complex. As genetic information on multi-factorial diseases is currently not used in the clinic, new protocols will have to be developed in order to guide health risk assessment for complex diseases. Especially since clinicians will have to deal with knowledge on many variants, all with minor effects on disease risk. A question that has to be answered is: should everybody´s genome be screened for any variant known, or should screening only be used to assess an individual´s risk in specific cases? Furthermore, the information flow produced by all the new studies that will be published needs to be structured in order to make it possible to use the knowledge into the clinic. Conclusively we can say that a lot of knowledge will be gained on the genetics of many diseases, but in order to use this information for health risk assessment in the clinic, many steps have to be taken.

5. Discussion

More than ten years ago, the first two surveys of the human genome were published. This was the beginning of a new era; the post-genome era. From the moment Collins and Venter were welcomed at the White House, people started to postulate expectations on what was about to change. One decade later we can look back on these expectations to see what has become reality and what has not. But even more important, we can look forward and anticipate on the changes that are about to 'move from bench to bed'. This thesis has presented an overview of the current state of personalized medicine and of the research that is being done in this new field. In this final chapter, several issues will be discussed, both practical and ethical, that researchers and clinicians will come across when personalized medicine is being implemented into clinical practice.

5.1 Conclusions

In current practice, personalized medicine is limited to several single non-genetic predictive factors. These have been determined in clinical trials that were done in order to distinguish two characteristics that influence treatment outcome. The example on breast cancer that is discussed in chapter 3, showed that both macroscopic and microscopic tumor characteristics can be used to determine the best type of primary and adjuvant treatment in cancer. When it comes to decision making in treatment, personal factors like age and gender are currently used as well. Risk predictions are used to determine if preventive measures have to be taken, or as a support for diagnosis. These predictions are based on family history, personal factors, biomarkers and environmental factors. Whole-genome screening for variants that cause disease or that are associated with elevated risk can currently only be done at commercial companies that offer a direct-to-customer SNP analysis. In many countries newborns are being screened, but only for a limited number of diseases of which the outcome can be influenced by early treatment. We can conclude that at this moment no systematic screening is done to determine a person's drug response, carrier statuses or disease risk in a whole genome fashion.

We have seen that there is a need for factors that have predictive value on drug response in order to avoid severe side effects and non-responders. Since the publication of the human genome, it is possible to screen for genetic variants that are responsible for differences in drug response among patients. This has led to the development of the field of pharmacogenomics. Especially the variants in ADME genes are currently investigated to determine their influence on the pharmacokinetics of drugs. Genome-wide association studies are done as well, through which genes and variants can be

determined that cause inter-patient variation both in pharmacokinetic and dynamic processes. In the future many drugs will be administered and prescribed to patients that, based on their genome, are likely to respond to drugs without developing severe side effects. In cancer, research is pointed towards the elucidation of prognostic factors. These factors are found by both genomics approaches (for example expression profiles) or other approaches like screening for CTCs.

Risk prediction will in the future be based on both genetic and non-genetic factors. Genetic factors are being determined through genome-wide association studies. The GWASs that are currently published mostly determined variants that have a small influence on the risk to develop complex diseases. For most Mendelian diseases the genes that are mutated have already been determined in the pre-genomic era, but as whole genome sequencing might be done for everyone in the future, not only patients but also carriers will come to light.

5.2 Implementation of pharmacogenomics

The research that has been done in the field of pharmacogenomics has led to optimism among many people. Prescription of pharmaceuticals based on response prediction has the potential to lead to higher drug efficacy and higher drug safety. Although these are promising results, there are still several hurdles that have to be overcome before pharmacogenomics can be incorporated into the clinic at large scale. The key players in the movement of pharmacogenomics from bench to bedside are the pharmaceutical industry, regulators, health care providers, physicians and patients.[70] For all of these, the main problems that will have to be solved are discussed here.

The pharmaceutical industry is a powerful industry, for which pharmacogenomics needs to be profitable in order for it to be incorporated into their products. Whether personalized medicine will show a return on investment (ROI) for pharmaceutical companies is still under debate[71]. Currently, companies are developing 'one size fits all' drugs. From a commercial, rather than a clinical perspective this is a profitable approach, since the products are used by a large spectrum of patients. Once drugs will only be prescribed to a selected group of patients, the investments made on the development of one product will have to be returned through selling fewer products. Therefore, either the prices of drugs will have to be raised or the prices for the development of drugs will have to be reduced. The clinical trials that are currently performed have to consist of many patients and controls and are therefore very expensive.[72] When patients can be assigned to treatment arms based on the presence or absence of a genetic variant that is being determined in conjunction with a new pharmaceutical, fewer patients will be required. Furthermore, the beneficial effects of the new

pharmaceutical will be more apparent in the targeted population and will not be obscured by a large number of non-responders.[73] Moreover, the introduction of pharmacogenomics brings a new opportunity for the industry to re-invest pharmaceuticals that were developed but not registered due to a lack of efficacy or safety. Many of the drugs that are discarded during any of the three phases of clinical trials, did work in a minor group of patients. Determination of possible genetic variants that caused this inter-patient variation, might lead to approval of the pharmaceutical with a pharmacogenomic label. Therefore it is possible that companies can get a ROI on these products, in which they have already invested money, while currently they would directly be discarded.

Regulators will play an important role in the incorporation of pharmacogenomics as well. As we have already seen in chapter 3, the US Food and Drugs Administration has already made it possible to administer drugs with a pharmacogenomic label. There is another vital role for regulators in this process. Currently, clinicians try to make informed decisions for patients, based on the best information available. As it is likely that prices of patient care will rise, regulators are responsible for keeping the best decision an affordable decision for the community. This can be done through the issuing of rules for the industry as well as for physicians. Furthermore, it is likely that, for the community, the costs of care will decrease since the society already bears a considerable burden for prescribed medication without benefit or eve for complications due to side effects.[73] Insurance companies will play a role as regulator as well. As these companies benefit from low health care costs, they may only reimburse pharmaceuticals that are prescribed based on genetic variants. In this way they will only have to pay for medication that will work, saving on medication that does not perform its function in a group of patients. If insurance companies include such clauses in their policies, this may be an impulse for pharmaceutical companies to develop medication in conjunction with a pharmacogenomic label, as non-label drugs might not be insured.

For physicians the economical aspects of the incorporation of personalized medicine into the clinic may only be a minor issue, since their primary goal is to give patients the best possible treatment. The major issue at this moment will be the lack of education of health professionals on topics like genomics as a diagnostic tool, genetic variation and molecular mechanisms of pharmacokinetics.[70] Appropriate education of physicians is essential for creating technical knowledge and awareness, that is needed in order to incorporate personalized medicine into the field of clinical medicine. Not only should courses on pharmacogenomics be taught in education programs in universities, all physicians that are currently working in medical practice need to be educated on pharmacogenomics. As in the field of pharmacogenomics new data are being generated with high speed, it will not be possible for each individual physician to keep updated on new information by

himself. Therefore it may be necessary to establish institutions that provide the relevant information in a summarized and ready-to-use way. Finally, the education of health care professionals is essential, since they have to be able to inform patients on the decisions that are made for his or her individual treatment.

5.2 Genome wide association studies

In chapter 4, the investigation of risk factors in complex diseases was discussed. Currently only non-genetic factors, like age or exposure to environmental factors, are being used for risk prediction. As risk prediction already plays a role in diagnostics and (preventive) treatment, one would expect genetic factors to be welcomed to complement the non-genetic factors. However, there is a lot of skepticism when it comes to GWASs. This has to do with several drawbacks of genome-wide association studies. One of the drawbacks is the fact that SNPs in both coding and non coding parts of the genome are being used. This means that very often, SNPs in non-coding regions of the genome are being found to be associated to a disease. It is hard to determine whether these SNPs are causal variants or whether they relate a genomic region to the disease. This makes it difficult to relate the SNP to the etiology of the disease.[73] Another drawback of the current application of GWASs, is that they are being used to determine variation between controls and patients that were clinically diagnosed. Disease classification is often based on a clinicians interpretation and a shared set of macro- and microscopic symptoms. Genetically though, they may be different diseases. This means that we do not actually know whether we are genetically (or etiologically) searching for the same disease. If this is not the case, a GWASs will not be able to determine variants associated with such a spectrum of diseases. One solution for the last problem could be to try to determine variants responsible for each symptom individually, as these might have a more homogenous etiology.

Skepticism will not keep the genome-wide association study away from the public, because as we have seen, direct-to-customer services already relate information from GWASs to an individual's genome sequence. Sooner or later clinicians will be confronted with a patient that has medical questions about the result of such a SNP analysis. Therefore the medical field should take the lead and again educate physicians in order to create health care professionals that are able to evaluate GWASs and their outcome. As the information flow that will originate from all the studies that will be published is too large to keep track of, consensus on GWASs will have to be created. The (bio)medical community should determine quality criteria for results of genome-wide association studies, that have to be met in order to be used in the clinic. Factors as study design, significance and minimal odds ratio for a risk variant (or a combination of variants) should be standardized. Once that is done,

new results can be centrally judged (by a responsible institution) where after they can be provided to and used by clinicians.

As was discussed in chapter 4, most variants that are currently associated with disease account for a small increase in risk. Therefore new methods will have to be used in order elucidate more of these variants and clarify the so called 'missing heritability'. Currently, the way genetic variants are linked to disease is by determining for each single variant whether or not it is associated with an increased risk to develop a certain disease. Once computational power increases and numerical algorithms are improved, it might be possible to look at combinations of variants rather than only single variants. This is useful, since the presence of a combination of for example two variants could be responsible for a significant increase in risk, whereas the two variants on their own are not associated with an increased risk. Furthermore, in most GWASs only common variants are studied. Although much effort is needed in order to study a population of patients with rare variants, the inclusion of rare variants into the GWAS, will probably lead to higher odds ratio's. Another advantage of rare variants is that they are (almost) always causal and will provide new insights in the etiology of disease. [74]

5.3 Ethical concerns in personalized medicine

The ethical discussions that have to be held in order to determine how personalized medicine should be incorporated into the clinic are still ongoing. As these discussions are essential for movement from bench to bedside, the topics will be touched upon in this section. These topics can basically be divided into three categories; 1. Privacy, 2. Sharing information with patients and 3. Obligations to close genetic relatives. Since the discussion on privacy is mainly a technical discussion on the protection of genetic information, this discussion will not be held here. I will focus on the privacy in relation to close genetic relatives.[75]

Anticipating the era of the $1000 genome is approaching, it will soon be possible to sequence the genome of each patient. Within a few decades it could even be cheaper to sequence the whole genome of a patient than to perform several genetic diagnostic tests during a patient's lifetime. Once someone's genomic sequence is available, the question raises who should be provided with this information and what information should be shared. Should this only be the validated data of clinical relevance, or should the patient or clinician get a file with 3 billion A's, G's, T's and C's? When a patient would be able to get all clinical relevant information, would any physician be obliged to act to all conclusions that could be drawn based on the analysis? Although currently these questions might seem irrelevant to many people, some clinicians are already dealing with them, since direct-to-

customer services do offer this kind information. Furthermore, clinicians need to know when a specific risk variant should be analyzed. This can be done at the moment of sequencing, in case of a positive family history, or if a patient is scared to develop the disease or only to support a diagnosis. It is clear that, at least at a national level, consensus among the medical society could guide medical decision making. Another topic that has to do with sharing of information is carrier status. Should everybody be screened for carrier statuses of Mendelian diseases? The question that follows has even more ethical implications: should a couple of two carriers be provided with the possibility to select an embryo that does not have both risk alleles? These questions are not easy to answer and it is well possible that consensus will never be achieved. In that case either (inter)national regulations or a physicians individual vision will determine what is appropriate in which situation.

Data that are obtained through genome sequencing do not only provide information on the person who is the source of the DNA, but also on his or her close relatives. Therefore, one could argue that in some cases information should be shared not only with the patient but also with close relatives. Even more controversial: should close relatives have equal say in whether or not some genetic analysis can be done? The general ethical issue is 'whether in the course of sequencing an individual's genome, someone is simultaneously obtaining identifiable private information about that person's close genetic relatives'[75]. Conclusively we can say that the ethical discussions should no longer be avoided, since personalized medicine is knocking on the door of medical practice.

This thesis has given an overview of the field of personalized medicine. By presenting and discussing various examples of research and clinical applications, many topics that play a role in personalized medicine have been highlighted. We can conclude that the expectations Collins postulated concerning the state personalized medicine in the year 2010 might have been too optimistic. Although research in the field of personalized medicine has led to the development of genetic tests, primary care providers are not yet starting to practice genetic medicine. As research is progressing, the clinicians need to be educated and involved in discussions on personalized medicine in order to create awareness and the knowledge that is necessary to make the step from bench to bedside. Whereas technical problems will have to be overcome by the field of research, the ethical and economical issues will have to be faced by the whole community in order to discuss these and create consensus on how the implementation of personalized medicine should take place. To my opinion these discussions will have to start now, since they will otherwise be overtaken by commercially available personalized medicine before the education of clinicians is completed and the ethical issues have led to consensus. Although it is hard to predict in what time span, the field of personalized medicine is very promising and will therefore change the way medicine is currently practiced.

6. Acknowlegdements

I would like to start by thanking all the people that participated enthusiastically in all the discussions on personalized medicine, during the writing of my thesis. Furthermore, I would like to thank Edwin Cuppen for providing access to 23andme. Finally I want to thank Wouter for the encouraging discussions and ofcourse for supervising me.

7. References

1. The human genome at ten. *Nature*. 2010;464:649-650.

2. Ginsburg GS, McCarthy JJ. Personalized medicine: Revolutionizing drug discovery and patient care. *Trends Biotechnol*. 2001;19:491-496.

3. Collins F. Has the revolution arrived? *Nature*. 2010;464:674-675.

4. Venter JC. Multiple personal genomes await. *Nature*. 2010;464:676-677.

5. Reiss T. Drug discovery of the future: The implications of the human genome project. *Trends Biotechnol*. 2001;19:496-499.

6. Metzker ML. Sequencing technologies the next generation. *Nature Reviews Genetics*. 2010;11:31-46.

7. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA: a cancer journal for clinicians*. 2010;60:277-300.

8. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thürlimann B, Senn H-. Thresholds for therapies: Highlights of the st gallen international expert consensus on the primary therapy of early breast cancer 2009. *Annals of Oncology*. 2009;20:1319-1329.

9. Gebauer G, Fehm T, Lang N, Jäger W. Tumor size, axillary lymph node status and steroid receptor expression in breast cancer: Prognostic relevance 5 years after surgery. *Breast Cancer Res Treat*. 2002;75:167-173.

10. Dunnwald LK, Rossing MA, Li CI. Hormone receptor status, tumor characteristics, and prognosis: A prospective cohort of breast cancer patients. *Breast Cancer Research*. 2007;9.

11. Morales L, Canney P, Dyczka J, et al. Postoperative adjuvant chemotherapy followed by adjuvant tamoxifen versus nil for patients with operable breast cancer: A randomised phase III trial of the european organisation for research and treatment of cancer breast group. *Eur J Cancer*. 2007;43:331-340.

12. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against her2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med*. 2001;344:783-792.

13. West M, Blanchette C, Dressman H, et al. Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci U S A*. 2001;98:11462-11467.

14. Javaher P, Nyoungui E, Kääriäinen H, et al. Genetic screening in europe. *Public Health Genomics*. 2010;13:524-537.

15. Larsson A. Neonatal screening for metabolic, endocrine, infectious, and genetic disorders: Current and future directions. *Clin Perinatol*. 2001;28:449-461.

16. Rowe SM, Clancy JP. Advances in cystic fibrosis therapies. *Curr Opin Pediatr*. 2006;18:604-613.

17. Farrell PM, Lai HJ, Li Z, et al. Evidence on improved outcomes with early diagnosis of cystic fibrosis through neonatal screening: Enough is enough! *J Pediatr*. 2005;147:S30-S36.

18. Castellani C, Southern KW, Brownlee K, et al. European best practice guidelines for cystic fibrosis neonatal screening. *Journal of Cystic Fibrosis*. 2009;8:153-173.

19. The international HapMap project: The international HapMap consortium. *Nature*. 2003;426:789-796.

20. Suh Y, Vijg J. SNP discovery in associating genetic variation with human disease phenotypes. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*. 2005;573:41-53.

21. Kiemeney LA, Thorlacius S, Sulem P, et al. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. *Nat Genet*. 2008;40:1307-1312.

22. Kiemeney LA, Sulem P, Besenbacher S, et al. A sequence variant at 4p16.3 confers susceptibility to urinary bladder cancer. *Nat Genet*. 2010;42:415-419.

23. Stern MC, Van Den Berg D, Yuan J-, et al. Sequence variant on 3q28 and urinary bladder cancer risk: Findings from the los angeles-shanghai bladder case-control study. *Cancer Epidemiology Biomarkers and Prevention*. 2009;18:3057-3061.

24. Wang M, Wang M, Zhang W, et al. Common genetic variants on 8q24 contribute to susceptibility to bladder cancer in a chinese population. *Carcinogenesis*. 2009;30:991-996.

25. Kaneko H, Kondo N. Clinical features of bloom syndrome and function of the causative gene, BLM helicase. *Expert Review of Molecular Diagnostics*. 2004;4:393-401.

26. Zhou K, Bellenguez C, Spencer CCA, et al. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nat Genet*. 2010.

27. Yoshiura K-, Kinoshita A, Ishida T, et al. A SNP in the ABCC11 gene is the determinant of human earwax type. *Nat Genet*. 2006;38:324-330.

28. Ginsburg GS, Willard HF. Genomic and personalized medicine: Foundations and applications. *Translational Research*. 2009;154:277-287.

29. Beitelshees AL, McLeod HL. Applying pharmacogenomics to enhance the use of biomarkers for drug effect and drug safety. *Trends Pharmacol Sci*. 2006;27:498-502.

30. Grossman I. ADME pharmacogenetics: Current practices and future outlook. *Expert Opinion on Drug Metabolism and Toxicology*. 2009;5:449-462.

31. Pasanen MK, Neuvonen PJ, Niemi M. Global analysis of genetic variation in SLCO1B1. *Pharmacogenomics*. 2008;9:19-33.

32. Kirchheiner J, Brockmöller J. Clinical consequences of cytochrome P450 2C9 polymorphisms. *Clin Pharmacol Ther*. 2005;77:1-16.

33. Goodsaid F, Frueh FW. Implementing the U.S. FDA guidance on pharmacogenomic data submissions. *Environ Mol Mutagen*. 2007;48:354-358.

34. Takeuchi F, McGinnis R, Bourgeois S, et al. A genome-wide association study confirms VKORC1, CYP2C9, and CYP4F2 as principal genetic determinants of warfarin dose. *PLoS Genetics*. 2009;5.

35. James AH, Britt RP, Raskino CL, Thompson SG. Factors affecting the maintenance dose of warfarin. *J Clin Pathol*. 1992;45:704-706.

36. O'Reilly RA. Studies on the optical enantiomorphs of warfarin in man. *Clin Pharmacol Ther*. 1974;16:348-354.

37. Takahashi H, Echizen H. Pharmacogenetics of CYP2C9 and interindividual variability in anticoagulant response to warfarin. *Pharmacogenomics Journal*. 2003;3:202-214.

38. Wadelius M, Chen LY, Downes K, et al. Common VKORC1 and GGCX polymorphisms associated with warfarin dose. *Pharmacogenomics Journal*. 2005;5:262-270.

39. Clarke M, Collins R, Davies C, Godwin J, Gray R, Peto R. Polychemotherapy for early breast cancer: An overview of the randomised trials. *Lancet*. 1998;352:930-942.

40. Van't Veer LJ, Dai H, Van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415:530-536.

41. Van De Vijver MJ, He YD, Van 'T Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*. 2002;347:1999-2009.

42. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst*. 2006;98:1183-1192.

43. Jacob K, Sollier C, Jabado N. Circulating tumor cells: Detection, molecular profiling and future prospects. *Expert Review of Proteomics*. 2007;4:741-756.

44. Xenidis N, Perraki M, Kafousi M, et al. Predictive and prognostic value of peripheral blood cytokeratin-19 mRNA-positive cells detected by real-time polymerase chain reaction in node-negative breast cancer patients. *Journal of Clinical Oncology*. 2006;24:3756-3762.

45. Camara O, Rengsberger M, Egbe A, et al. The relevance of circulating epithelial tumor cells (CETC) for therapy monitoring during neoadjuvant (primary systemic) chemotherapy in breast cancer. *Annals of Oncology*. 2007;18:1484-1492.

46. Mocellin S, Keilholz U, Rossi CR, Nitti D. Circulating tumor cells: The 'leukemic phase' of solid cancers. *Trends Mol Med*. 2006;12:130-139.

47. Grabinski JL. Pharmacogenomics of anticancer agents: Implications for clinical pharmacy practice. *Journal of Pharmacy Practice*. 2007;20:246-251.

48. De Jongh TOH, De Vries H, Grundmeijer HGLM. *Diagnostiek Van Alledaagse Klachten.* Vol 2. 2nd ed. Houten: Bohn Stafleu van Loghum; 2005.

49. Leander K, Hallqvist J, Reuterwall C, Ahlbom A, De Faire U. Family history of coronary heart disease, a strong risk factor for myocardial infarction interacting with other cardiovascular risk factors: Results from the stockholm heart epidemiology program (SHEEP). *Epidemiology*. 2001;12:215-221.

50. Lloyd-Jones DM, Nam B-, D'Agostino Sr. RB, et al. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-aged adults: A prospective study of parents and offspring. *J Am Med Assoc*. 2004;291:2204-2211.

51. Zlot AI, Valdez R, Han Y, Silvey K, Leman RF. Influence of family history of cardiovascular disease on clinicians' preventive recommendations and subsequent adherence of patients without cardiovascular disease. *Public Health Genomics*. 2010;13:457-466.

52. Bello N, Mosca L. Epidemiology of coronary heart disease in women. *Prog Cardiovasc Dis*. 2004;46:287-295.

53. Molenaar EA, Hwang S-, Vasan RS, et al. Burden and rates of treatment and contr of cardiovascular disease risk factors in obesity: The framingham heart study. *Diabetes Care*. 2008;31:1367-1372.

54. Yusuf S, Reddy S, Ôunpuu S, Anand S. Global burden of cardiovascular diseases. part I: General considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation*. 2001;104:2746-2753.

55. Bittner V. Menopause, age, and cardiovascular risk. A complex relationship. *J Am Coll Cardiol*. 2009;54:2374-2375.

56. Arsenault BJ, Boekholdt SM, Kastelein JJP. Lipid parameters for measuring risk of cardiovascular disease. *Nature Reviews Cardiology*. 2011.

57. Koukkunen H, Penttilä K, Kemppainen A, et al. Troponin T and creatinine kinase isoenzyme MB mass in the diagnosis of myocardial infarction. *Ann Med*. 1998;30:488-496.

58. Insel PA, Tang C-, Hahntow I, Michel MC. Impact of GPCRs in clinical medicine: Monogenic diseases, genetic variants and drug targets. *Biochimica et Biophysica Acta - Biomembranes*. 2007;1768:994-1005.

59. Botstein D, Risch N. Discovering genotypes underlying human phenotypes: Past successes for mendelian disease, future approaches for complex disease. *Nat Genet*. 2003;33:228-237.

60. Jmoudiak M, Futerman AH. Gaucher disease: Pathological mechanisms and modern management. *Br J Haematol*. 2005;129:178-188.

61. Nieuwenhuis MH, Vasen HFA. Correlations between mutation site in APC and phenotype of familial adenomatous polyposis (FAP): A review of the literature. *Crit Rev Oncol*. 2007;61:153-161.

62. OMIM. Available at: http://www.ncbi.nlm.nih.gov/Omim/mimstats.html. Accessed 04/03, 2011.

63. Ku CS, Loy EY, Pawitan Y, Chia KS. The pursuit of genome-wide association studies: Where are we now. *J Hum Genet*. 2010;55:195-206.

64. Burton PR, Clayton DG, Cardon LR, et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661-678.

65. Satake W, Nakabayashi Y, Mizuta I, et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for parkinson's disease. *Nat Genet*. 2009;41:1303-1307.

66. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Current Opinion in Genetics and Development*. 2009;19:212-219.

67. Klein RJ, Zeiss C, Chew EY, et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005;308:385-389.

68. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. *J Clin Invest*. 2008;118:1590-1605.

69. Roberts R, Wells GA, Stewart AFR, Dandona S, Chen L. The genome-wide association study-A new era for common polygenic disorders. *Journal of Cardiovascular Translational Research*. 2010;3:173-182.

70. Freuh FW, Gurwitz D. From pharmacogenetics to personalized medicine: A vital need for educating health professionals and the community. *Pharmacogenomics*. 2004;5:571-571-579.

71. Keckley PH, Dhar A, Underwood HR. The ROI for targeted therapies: A strategic perspective. *Deloitte Center for Health Solutions*. 2009.

72. Eisenstein EL, Collins R, Cracknell BS, et al. Sensible approaches for reducing clinical trial costs. *Clinical Trials*. 2008;5:75-84.

73. Leeder JS, Spielberg SP. Personalized medicine: Reality and reality checks. *Ann Pharmacother*. 2009;43:963-966.

74. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet*. 2008;40:695-701.

75. McGuire AL, Caulfield T, Cho MK. Research ethics and the challenge of whole-genome sequencing. *Nature Reviews Genetics*. 2008;9:152-156.