Research Methods for Clinical Trials in Personalized Medicine: A Systematic Review

Zheng Ren*         Marie Davidian†         Stephen L. George‡
Richard M. Goldberg**  Fred A. Wright††
Anastasios A. Tsiatis‡‡  Michael R. Kosorok§

*University of North Carolina at Chapel Hill, renzheng@email.unc.edu
†North Carolina State University, davidian@ncsu.edu
‡Duke University, georg001@mc.duke.edu
**Ohio State University, richard.goldberg@osumc.edu
††University of North Carolina at Chapel Hill, fred_wright@unc.edu
‡‡North Carolina State University, tsiatis@stat.ncsu.edu
§University of North Carolina at Chapel Hill, kosorok@unc.edu

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

http://biostats.bepress.com/uncbiostat/art25

Copyright ©2012 by the authors.
Research Methods for Clinical Trials in Personalized Medicine: A Systematic Review

Zheng Ren, Marie Davidian, Stephen L. George, Richard M. Goldberg, Fred A. Wright, Anastasios A. Tsiatis, and Michael R. Kosorok

Abstract

Background: Personalized medicine, the notion that an individual’s genetic and other characteristics can be used to individualize the diagnosis, treatment and prevention of disease, is an active and exciting area of research, with tremendous potential to improve the health of society.

Methods: Seventy-six studies using personalized medicine analysis techniques published from 2006 to 2010 in six high-impact journals - Journal of the American Medical Association, Journal of the National Cancer Institute, Lancet, Nature, Nature Medicine, and the New England Journal of Medicine - were reviewed. Selected articles were manually selected based on reporting of the use of genetic information to stratify subjects and on analyses of the association between biomarkers and patient clinical outcomes.

Results: We found considerable variability and limited consensus in approaches. Approaches could largely be classified as data-driven, seeking discovery through statistical analysis of data, or knowledge-driven, relying heavily on prior biological information. Some studies took a hybrid approach. Eliminating two articles that were retracted after publication, 56 of the remaining 74 (76%) were cancer-related.

Conclusions: Much work is needed to standardize and improve statistical methods for finding biomarkers, validating results, and efficiently optimizing better individual treatment strategies. Several promising new analytic approaches are available and should be considered in future studies of personalized medicine.
Research Methods for Clinical Trials in Personalized Medicine: A Systematic Review

Zheng Ren, Marie Davidian, Stephen L. George, Richard M. Goldberg, Fred A. Wright, Anastasios A. Tsiatis, Michael R. Kosorok

Affiliations of authors: Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC (ZR, FAW, MRK); Department of Statistics, North Carolina State University, Raleigh, NC (MD, AAT); Cancer Center Biostatistics, Duke University Medical Center, Durham, NC (SLG); Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC (SLG, MD, AAT); Comprehensive Cancer Center, The Ohio State University Medical Center, Columbus, OH (RMG); Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, Chapel Hill, NC (MRK).

Correspondence to: Micheal R. Kosorok, PhD, Department of Biostatistics and Department of Statistics and Operations Research, University of North Carolina-Chapel Hill, 3101 McGavran-Greenberg Hall, CB 7420, Chapel Hill, NC 27599-7420 (e-mail: kosorok@unc.edu).
Abstract

Background: Personalized medicine, the notion that an individual’s genetic and other characteristics can be used to individualize the diagnosis, treatment and prevention of disease, is an active and exciting area of research, with tremendous potential to improve the health of society.

Methods: Seventy-six studies using personalized medicine analysis techniques published from 2006 to 2010 in six high-impact journals - Journal of the American Medical Association, Journal of the National Cancer Institute, Lancet, Nature, Nature Medicine, and the New England Journal of Medicine - were reviewed. Selected articles were manually selected based on reporting of the use of genetic information to stratify subjects and on analyses of the association between biomarkers and patient clinical outcomes.

Results: We found considerable variability and limited consensus in approaches. Approaches could largely be classified as data-driven, seeking discovery through statistical analysis of data, or knowledge-driven, relying heavily on prior biological information. Some studies took a hybrid approach. Eliminating two articles that were retracted after publication, 56 of the remaining 74 (76%) were cancer-related.

Conclusions: Much work is needed to standardize and improve statistical methods for finding biomarkers, validating results, and efficiently optimizing better individual treatment strategies. Several promising new analytic approaches are available and should be considered in future studies of personalized medicine.
The emerging field of personalized medicine has been the focus of vigorous research in recent years. Since the initiation of the Human Genome Project twenty years ago (1), researchers have successfully identified tens of thousands of genes and studied their protein products and functions. More recently, genetic typing and sequencing techniques have revolutionized our understanding of individual genetic variation (2,3), spurring research on new therapeutic agents that can target genetic differences across individuals. Differences in therapeutic response among individuals may arise from a number of other sources, such as cumulative environmental exposures, which may be difficult to assess. The genetic revolution has thus provided a measurable basis for personalized medicine, and has placed the field in the national spotlight. The goal of the Genomics and Personalized Medicine Act of 2010 (GPMA), introduced in the United States House of Representatives in May 2010, is to encourage genomics research and the application of personalized medicine (4); a similar act was introduced previously in the United States Senate by then-Senator Barack Obama. As public awareness and interest in genetic testing continues to heighten, research on the use of genomic and other patient information to improve the prevention, diagnosis and treatment of disease is likely to accelerate.

The ultimate goal of personalized medicine is to make the optimal treatment decision based on all of the information from a patient, including not only genetic information but also age, gender, other clinical factors, current disease status and other information. Methods for exploiting very large quantities of data, integrated with biological knowledge, and innovative study designs are critical to this goal. In this article, we review approaches that have been applied in published studies to use patient genomic
information as the basis for improving the efficacy of treatments and reducing toxicity. Issues and challenges in the design and analysis of personalized medicine studies are also discussed, and new approaches are proposed.

**Definition of Personalized Medicine**

There is no single definition of “personalized medicine,” and differing terminology has been used in the literature by various stakeholders to describe the concept, including personalized medicine (5-8), personalized therapy (9,10), personalized drugs (11) and tailored therapy (12,13). The National Institutes of Health (NIH), for example, defines personalized medicine as “an emerging practice of medicine that uses an individual's genetic profile to guide decisions made in regard to the prevention, diagnosis, and treatment of disease. Knowledge of a patient's genetic profile can help doctors select the proper medication or therapy and administer it using the proper dose or regimen.” For the U.S. Food and Drug Administration (FDA), the goal of personalized medicine is “to get the best medical outcomes by choosing treatments that work well with a person's genomic profile or with certain characteristics in the person's blood proteins or cell surface proteins” (14). In the GPMA, personalized medicine is defined as “any clinical practice model that emphasizes the systematic use of preventive, diagnostic, and therapeutic interventions that use genome and family history information to improve health outcomes.” Recently, Margaret A. Hamburg, Commissioner of the FDA, and Francis S. Collins, Director of NIH, shared their joint opinion about the concept (15), proposing that the goal of personalized medicine is to identify the best treatment strategy to achieve the optimal clinical outcome based on a patient’s individual profile. Note that the definitions encompass variation in constitutional genetic profile, which is stable for
an individual, as well as more variable genomic or other biomarkers that may be specific to a disease state, tumor, or tissue.

**The Promise of Personalized Medicine**

Personalized medicine has demonstrated the potential to lead to enhanced treatment effect in clinical studies of already-approved therapies within gene-profile stratified subgroups (16), to avoid unnecessary toxicity associated with treatments (17) and to minimize the rate of adverse events (18). In many settings, it is likely that some subgroups of patients will have a better response to a new treatment and others will have a worse response. If the population is considered as a whole, the overall difference would be attenuated by those with poorer responses, and no significant improvement would be detected for the population (Figure 1). Therefore, personalized medicine has the potential to improve the performance of therapy for individual patients through recognizing and capitalizing on patient heterogeneity. Note that the very concept of personalized medicine represents a profound shift in the interpretation of treatment data, especially in the context of controlled clinical trials, for which subgroup analysis has often been viewed with suspicion as a potential source of spurious findings.

A striking example of meaningful subgroup analysis occurred in a randomized prospective phase III clinical trial in patients with advanced non-small cell lung cancer that confirmed the benefit of first-line gefitinib, an epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor, to chemotherapy. In subjects with EGFR mutation-positive tumors, PFS was significantly improved for gefitinib versus chemotherapy (median PFS, 9.5 months vs. 6.3 months), suggesting that the presence of EGFR positive
mutation should be used as a criterion for determining first-line therapy for the patients with advanced non-small-cell lung cancer (19).

**Methods**

**Literature Search**

In this review, we summarize studies related to personalized medicine that were published in six high-impact journals - Journal of the American Medical Association, Journal of the National Cancer Institute, Lancet, Nature, Nature Medicine, and the New England Journal of Medicine - from 2006 to 2010. All papers were manually selected and reviewed. The inclusion criteria were: 1) papers published between January 1, 2006 and December 31, 2010; 2) papers reporting the results of original clinical studies using genetic information obtained from DNA, RNA, protein or tumor mutations to stratify the study population; and 3) papers reporting the association between biomarkers and patient clinical outcomes. Exclusion criteria were: 1) papers focused mainly on the association between genetic profile and incident disease; 2) papers focused on identifying biomarkers for disease diagnosis; and 3) papers reporting systematic reviews and meta-analysis. Seventy-six articles (12 papers published in Journal of the American Medical Association, 17 papers published in Journal of the National Cancer Institute, 9 papers published in Lancet, 6 papers published in Nature, 7 papers published in Nature Medicine, and 25 papers published in the New England Journal of Medicine) were selected meeting the above criteria (16,20-94). As discussed shortly, two of these (74,87) have since been retracted and are not included in the summaries in Table 1 or elsewhere in this article. The predominance of cancer studies reflects the fact that cancer inherently involves genetic dysregulation, increasing the likelihood that suitable biomarkers and
targeted therapies may be identified. The genomic data in most of the reviewed studies consisted of transcriptional expression, profiled by microarray or polymerase chain reaction (PCR).

**Results**

**Approaches to Finding Biomarkers for Personalized Medicine**

A complete process leading to effective personalized medicine will typically include the following five key elements: obtaining patient genetic/genomic data using array and other high throughput technology (95); identifying one or more biomarkers (as discussed shortly) (96); developing new or selecting available therapies (97); measuring the relationship between biomarkers and clinical outcomes, including the prognosis and response to therapy; and verifying the relationship in a prospective randomized clinical trial (98).

A biomarker is an indicator that can potentially stratify patients into subgroups with common features. Biomarkers can consist of DNA, RNA, proteins and types of tumors. Because of the so-called “curse of dimensionality” (99), where there are vast numbers of genes relative to patient samples, identifying a potential biomarker from patient genomic profiles is a tremendous challenge for geneticists and statisticians. Many techniques for the development and validation of biomarkers are available (100,101). In the studies reviewed, two main approaches were taken: a data-driven approach (102) and a knowledge-driven approach (103) (Figure. 2).

In the data-driven approach, extensive genetic information is collected on each participant, and an analysis is conducted to select those genes or sets of genes that are most strongly associated with the targeted clinical outcome. The strength of this approach
is that it is unbiased in the sense that all genes are considered and their relationship to clinical outcome is tested. During the selection process, new genes or gene functions can potentially be identified as involved in disease. For example, in one simple data-driven approach, information on response to a therapy and gene expression data is collected on each patient. Each gene’s expression is compared between the response and non-response groups, and p-values are calculated using t-tests; then the p-value for each gene is adjusted upward to account for multiple comparisons via the Bonferroni method and based on the number of genes used in the tests (104). Genes with significant adjusted p-values are considered to be potential biomarkers for the response to the therapy, which can be verified in follow-up trials (35).

A typical data-driven approach attempts to identify a potential biomarker profile constructed from one or more genes from among tens of thousands, a process complicated by the correlation structure and multiplicity of high dimensional data. Many statistical methods have been used to identify promising genetic patterns from numerous genes, including the correlation coefficient (12), t-tests (82), hierarchical clustering (105), Significance Analysis of Microarrays (SAM) (106), Prediction Analysis for Microarrays (PAM) (107), and principal component analysis (PCA) (108). Once the number of candidate genes is narrowed down to several hundred, other statistical techniques can be applied, such as the proportional hazards model (109) or a Receiver Operator Characteristic (ROC) (110) analysis. In one microarray study, the expression of 25,000 genes was measured in 78 lymph-node-negative breast cancers, and 5,000 genes expressed in more than 3 of 78 tumors were selected. The correlation coefficient of the expression for each gene with disease outcome was calculated and the 5,000 genes were
further narrowed to 231 (correlation coefficient <-0.3 or >0.3). This group of 231 genes was rank-ordered on the basis of the magnitude of the correlation coefficient. A subset of 5 genes was subsequently added to the prognosis classifier until the accuracy was maximized. Seventy genes were identified to constitute a signature based on correlation coefficients for predicting clinical outcome of breast cancer (12). The signature was validated in a follow up study, and it outperformed the current standard method in predicting all endpoints (37).

Data quality is a key issue for the data-driven approach and, indeed, for any approach. Several studies with similar experimental designs generated different gene signatures from similar populations based on this approach, possibly owing to varying data quality (103,111,112). Moreover, gene signatures identified by this approach are difficult to interpret because the genes may not be causal but may result instead from correlation with some underlying causal mechanism. In addition, large sets of genes are highly correlated, and additional follow-up studies are usually required to refine the gene signature. New statistical methods and software continue to be proposed to identify gene signatures in a more efficient and precise way. Several articles have systematically reviewed such statistical methods (103,113,114), but there is no clear consensus in the statistics and data mining community about which methods should be used for data-driven gene signature identification.

The knowledge-driven approach is based on existing knowledge of the biological functions of certain genes or proteins that may arise, for example, through a pathway analysis to establish a mechanism for growth inhibition in tumor cells (8). A targeted inhibitor that may benefit patients with a specific gene mutation can be designed using
this approach, as with cetuximab and panitumumab, approved by the FDA and targeting the epidermal growth factor receptor in multiple cancers (115). The knowledge-driven approach can also help limit the number of candidate genes investigated. For example, only pre-determined single nucleotide polymorphisms (SNPs) or proteins might be considered as biomarkers. This was illustrated in a recent metastatic colorectal cancer study, where 15 candidate genes were investigated as predictors of adverse events and response to chemotherapy. Finally, four genes showed significant association with adverse events (116). This method avoids the issues associated with interpreting high-dimensional data that arise in the data-driven approach and improves the accuracy of the analysis; however, the drawback is that the results are only as good as the existing state of knowledge: genes that are not already known to be involved in a process are not considered.

Some studies in our sample combined the two approaches to develop a gene signature. The data-driven approach is employed first to reduce the number of available genes from tens of thousands to several hundred, and then the knowledge-driven approach employs a pathway analysis. Only the genes that appear on both lists are included in the final gene signature. Grouping genes according to different biological pathways is another way to combine the two approaches and reduce the dimension of available information efficiently.

Overall, 56 of 74 studies in our sample explored new biomarkers and eighteen studies validated the biomarkers discovered in previous studies. Among fifty-six exploratory studies, sixteen of them used primarily a data-driven approach, thirty-six studies used a knowledge-driven approach, and four studies used a combined approach. Across the 74
studies, 62 used a gene signature comprising fewer than ten genes, mutations or SNPs as the biomarker to stratify the study population, perhaps because biomarkers with fewer genes are easier to analyze and interpret.

**Prognostic and Predictive Biomarkers**

In the studies reviewed, two types of relationships between biomarkers and clinical outcomes were investigated, association between biomarkers and patient prognosis and association between biomarkers and response to treatment. In common parlance, the first type of analysis refers to “prognostic” biomarkers and the second to “predictive” biomarkers. In the study of potential prognostic biomarkers, patient information is pooled and a prognostic indicator is generated based on patient clinical outcomes without considering the different treatments or perhaps restricting to some standard treatment. The prognostic indicator is expected to predict patient survival without regard to the treatments the patients received. A good example would be the 70-gene signature for breast cancers mentioned above. This kind of investigation can be conducted by pooling available data sets from several similar studies. The results can help clinicians predict patient performance but cannot be used to assign the proper treatment to patients based on their genetic profiles and clinical factors. In the study of potential predictive biomarkers, responses to different treatments are compared within biomarker strata and between patients with different biomarkers. These studies require well-designed clinical trials to verify the hypotheses and to show efficacy in and between biomarker groups. In the study conducted by Amado et al., the effectiveness of panitumumab was compared between patients with wild type and mutations of KRAS in colorectal cancers in a phase
III clinical trial. Panitumumab was shown to significantly improve the survival of patients with wild type KRAS as compared to best supportive care alone, but there was no significant improvement for mutation carriers. KRAS mutations showed resistance to panitumumab treatment, and this finding was confirmed in several follow-up studies (116). Results obtained from both types of investigations are important for clinical practice, but knowledge that a specific treatment elicits different outcomes across subgroups is more valuable for personalized medicine, because it can guide an optimal treatment plan based on biomarker information. For example, it has been shown that patients with Human Epidermal Growth Factor Receptor 2 (HER2)-Positive Breast Cancer have a worse prognosis than other subgroups, but have a better response to trastuzumab, which is designed to interfere with the HER2 receptor (117). According to these results, a patient with HER2-positive breast cancer would be given trastuzumab to maximize treatment benefit.

In the reviewed literature, nineteen studies compared the efficacy of different treatments for one group of patients; thirty-three investigated the effect of the same therapy among different subgroup populations; and sixteen studies made both types of comparisons.

**Discussion**

**Reliability and Reproducibility of Personalized Medicine**

A continuing controversy of personalized medicine focuses on its reliability and reproducibility. Some personalized medicine studies have been reported as non-replicable, and the related clinical trials have been suspended and scrutinized (118,119); indeed, two of the papers reviewed were retracted on this basis (74,87). The complexity
of the data and statistical analyses involved in personalized medicine make study of reproducibility of results essential. Hence the associated datasets must be made publicly available for verification of results. The obtained biomarkers need to be validated in a different group of patients to avoid potential selection bias from the study population. Because extensive, high-dimensional data are involved in many personalized medicine studies, data management is another important issue. Some recommendations for improved data management in genetics research and personalized medicine have been made (2,120-122).

**Future Clinical Trial Designs for Personalized Medicine**

The identification and verification of new biomarkers requires significant resources of time, money and patients. Developing efficient and effective methods for the discovery and verification of new biomarkers is a fundamental challenge. There are two general statistical approaches for assessing personalized treatment regimens in current work.

One general approach relies on standard regression models of clinical outcomes given treatment and other covariates. Interactions between treatment and covariates are included in such regression models, and the optimal decision rules are obtained by identifying the treatment leading to the largest mean clinical outcome based on the model for any combination of covariate values. Some new clinical trial designs for personalized medicine have been developed using this approach, such as the biomarker-adaptive threshold design (123-125) and Bayesian designs (126), including the currently active Investigation of Serial studies to Predict Your Therapeutic Response with Imaging And moLecular analysis (I-SPY 2) trial, which involves an adaptive design to investigate biomarkers for women with breast cancer (Figure. 3) (127). These designs allow
comparison of several different biomarkers with different treatments in the same clinical trial. Bayesian designs have also been implemented in ongoing phase III clinical trials to find the best treatment for patients with different biomarkers (126).

The second general approach uses machine learning techniques to determine the relationship between clinical outcomes and treatment plus characteristics such as age, gender and biomarkers (128). These methods, used extensively in artificial intelligence, lead to models for these relationships that are more flexible than those in standard regression analysis and thus can more easily describe complex relationships. Designs based on these techniques have been developed and implemented for chronic diseases such as cancer and have shown the capability of discovering the best treatment for each individual based on a reasonable sample size (129,130); however, they have not yet been implemented in practice. The machine learning technique is expected to discover new relationships between biomarkers and patient response to treatments without any prior knowledge about the true relationship, similar to what the data-driven approach does for identifying biomarkers.

It is important to recognize, especially in the case of chronic diseases and conditions, that “treatment” is often an ongoing process involving a series of decisions made over time. For example, the treatment of cancer may involve selection of a first line of chemotherapy and observation of the patient’s response, followed by selection of a second line therapy in the event of no response or of a maintenance or intensification treatment if a positive response is elicited. At each decision point, issues of selection of therapy among available options, choice of dosage and timing of therapy, and so on must be addressed. Identifying the overall strategy for making the entire sequence of decisions
to optimize clinical outcome, taking into account the evolving status of the patient at each point, would be a major advance for personalized medicine. It is possible that the treatment that appears “best” with respect to clinical outcome for a single decision point may not be the best choice when viewed in the context of an entire series of decisions. For example, a treatment given early may cause bone marrow toxicity that compromises the dose intensity of future cytotoxic regimens, thus rendering subsequent treatments less effective.

Methods for determining optimal decision-making strategies for personalized medicine using existing databases have been developed (131) but are limited by the quality of the data, as all information relevant to each decision point may not have been recorded. A promising alternative approach is to conduct a prospective clinical trial focused on this goal using the so-called Sequential Multiple Assignment Randomized Trial (SMART) design (132). SMARTs have been conducted for this purpose in the study of substance abuse and depression, and a similar design suitable for trials involving life-threatening diseases has been developed (129). In a typical SMART design, participants are initially randomized to treatment options based on their profile; at the end of the first treatment segment, patient response to the first treatment and status are reviewed and the second treatment decision is made based on response and status; randomization of treatment may also be employed at this step and later steps; and the steps are repeated until the end of study. The best decision strategy for each individual can be generated from such designs using artificial intelligence techniques such as reinforcement learning (122,123), as has been demonstrated for lung cancer. For example, this dynamic decision method can discover the best personalized therapy in a
complex clinical trial setting involving multiple treatments, multiple biomarkers and several decision points.

**Conclusions**

Personalized medicine has demonstrated considerable promise, although many additional challenges need to be addressed and interpreted in future studies. Even though biomarker and related clinical studies have shown potential application in clinical practice, there is no strong evidence that shows that gene-based biomarkers can completely replace traditional clinical characteristics (133); using the biomarkers in clinical trials can cause unbalanced randomization, which can produce biased assessment of treatment (134). In addition, gene-gene and gene-environment interactions potentially should also be included and modeled using high dimensional data analysis methods. Addressing all these challenges will require the development of new statistical tools for finding biomarkers, validating results and discovering better decision making processes for selecting the optimal treatment strategy in an efficient way. This endeavor will not be easy, but the potential gains in human health are enormous.

**References**


Table 1. Summary of personalized medicine studies published between 2006 and 2010 in six journals

| The number of genes or proteins in each biomarker | Greater than 100 genes | 4 (5%) |
| Sample size | Larger than 1000 patients | 17 (24%) |
| Cancer | Between 500 and 1000 patients | 12 (16%) |
| | between 100 and 500 patients | 35 (47%) |
| | Less than 100 | 10 (14%) |

| Cancer | Breast cancer (16,20,28,33, 35,37,38,41,43,45,52, 54,55,62,66,75,76,83-85,91,94) | 22 (30%) |
| Lung cancer (26,44,58,60,61,67,70,81,88) | 9 (12%) |
| Leukemia (22,31,39,46,65,69,71,80) | 8 (11%) |
| Colorectal cancer and adenomas (21,27,30,36,42,59,64) | 7 (9%) |
| Gastric cancer (49) | 1 (1%) |
| Liver Cancer (63) | 1 (1%) |
| Lymphomas (93) | 1 (1%) |
| Melanoma (48) | 1 (1%) |
| Mesothelioma (40) | 1 (1%) |
| Neuroblastomas (34) | 1 (1%) |
| Ovarian carcinomas (92) | 1 (1%) |
| Prostate cancer (47) | 1 (1%) |
| Sarcomas (82) | 1 (1%) |
| Skin cancer (75) | 1 (1%) |

| Non-Cancer | Cardiovascular (23,24,29,51,56,72,79,80) | 8 (11%) |
| Asthma (50,57) | 2 (3%) |
| Hepatitis C virus (89,90) | 2 (3%) |
| Thrombosis and anti-coagulation (32,78) | 2 (3%) |
| Autoimmune disease (86) | 1 (1%) |
| Diabetes (53) | 1 (1%) |
| HIV (68) | 1 (1%) |
| Kidney Transplant (25) | 1 (1%) |
**Figure Legends**

Figure 1:

The comparisons among subgroups.

Figure 2:

Two approaches for identifying genetic biomarkers for personalized medicine.

Figure 3:

The overall trial design for I-SPY 2: MRI and blood sample draw is used to determine biomarker signature and eligibility; patients are randomized to seven treatment groups including two control groups and five new treatment groups. After 3 weeks of the assigned treatment, patients receive a repeat MRI and core biopsy and finish treatment over 9 additional weeks. A third MRI and core biopsy are performed before initiating standard chemotherapy with doxorubicin and cyclophosphamide. Another blood sample draw and a fourth MRI are given prior to surgery. Tumor tissue is collected at surgery to assess whether the patient has pathologic complete response, which is used to recalculate the randomization probability of each treatment group for new patients.
Figure 1. Comparisons among subgroups.
Figure 2. Two approaches for identifying genetic biomarkers for personalized medicine.
Figure 3. The overall trial design for I-SPY 2: MRI and blood sample draw is used to determine biomarker signature and eligibility; patients are randomized to seven treatment groups including two control groups and five new treatment groups. After 3 weeks of the assigned treatment, patients receive a repeat MRI and core biopsy and finish treatment over 9 additional weeks. A third MRI and core biopsy are performed before initiating standard chemotherapy with doxorubicin and cyclophosphamide. Another blood sample draw and a fourth MRI are given prior to surgery. Tumor tissue is collected at surgery to assess whether the patient has pathologic complete response, which is used to recalculate the randomization probability of each treatment group for new patients.