예방의학회지 제42권 제6호(2009년 11월)
J Prev Med Public Health 2009;42(6):371-376
DOI: 10.3961/jpmph.2009.42.6.371

Special Issues

Risk Assessment and Pharmacogenetics in Molecular and Genomic Epidemiology

Sue K. Park[1,2], Ji-Yeob Choi[3]

Department of Preventive Medicine, Seoul National University College of Medicine[1]; Cancer Research Institute and Institute of Health Policy and Management, Seoul National University[2]; Pharmacogenomics Research Center, Inje University College of Medicine[3]

In this article, we reviewed the literature on risk assessment (RA) models with and without molecular genomic markers and the current utility of the markers in the pharmacogenetic field. Epidemiological risk assessment is applied using statistical models and equations established from current scientific knowledge of risk and disease. Several papers have reported that traditional RA tools have significant limitations in decision-making in management strategies for individuals as predictions of diseases and disease progression are inaccurate. Recently, the model added information on the genetic susceptibility factors that are expected to be most responsible for differences in individual risk. On the continuum of health care, from diagnosis to treatment, pharmacogenetics has been developed based on the accumulated knowledge of human genomic variation involving drug distribution and metabolism and the target of action, which has the potential to facilitate personalized medicine that can avoid therapeutic failure and serious side effects. There are many challenges for the applicability of genomic information in a clinical setting. Current uses of genetic markers for managing drug therapy and issues in the development of a valid biomarker in pharmacogenetics are discussed.

J Prev Med Public Health 2009;42(6):371-376

Key words : Risk assessment, Pharmacogenetics, Biologic markers, Genetic markers

INTRODUCTION

There is a growing interest in the epidemiologic research field of genomic molecular based health care for early detection of illness and individual therapy. Many common and complex diseases are generally considered to be a multifactorial disorders; understanding genetic and molecular data and the related multifactorial etiological factors is essential for risk assessment. The goal of the molecular and genomic epidemiologic field is to understand the effects of various exposures on disease outcome and the genetic variability that may alter the prevention and treatment of disease. Molecular genomic markers can play a role in more accurately predicting disease development, drug response, and dosage optimization of drugs by adding markers to previous risk assessment tools. This has implications for personalized predictive prevention of disease development and progression, personalized treatment of diseases, and drug discovery and development. In this article, we reviewed the literature related to risk assessment models with and without molecular genomic markers and the current utility of the markers in the pharmacogenetic area.

RISK ASSESSMENT IN MOLECULAR AND GENOMIC EPIDEMIOLOGY

Risk assessment (RA), a step in the risk management process in environmental and occupational health, estimates the probability of a hazard effect of risk factors in individuals or populations. Environmental and occupational health focuses on chemicals and toxic materials above a certain threshold that are risk factors for various diseases and death. In environmental and occupational health, the risk assessment process includes 'Hazard Identification', a 'Dose-response Assess- ment' between the exposure dose and probability or incidence of effect, and 'Exposure Quantification' for the determina-tion of the dosage that people receive [1].

In epidemiology, risk assessment estimates the chances of developing a disease over a specified interval in a person with certain risk factors. Epidemiological risk assessment is applied using statistical models and equations established from current scientific knowledge pertaining to etiology and the risk factors of disease. The RA models used in epidemiology are classified into four types according to the aim, study design of the baseline data, statistical model construction, and target population. These are summarized in Table 1.

The first is the absolute risk model. In this model, a cohort study with sufficient follow-up was used to compute the real probability of disease development in a cohort with sufficient follow-up periods. This model computes the probability of a person's risk using Cox's proportional hazard model. The disease

Received : 2 October 2009, Accepted : 2 November 2009
Corresponding author : Ji-Yeob Choi(633-165 Gaegum-dong, Jin-gu, Busan 614-735, Korea, Tel : +82-51-890-8663, Fax : +82-51-893-1232, E-mail : jychoi@inje.ac.kr)

Table 1. Classification of epidemiological risk assessment (RA) models according to aim, baseline data, target population, and statistical method

Type	Characteristics	Examples
Absolute RA model	Aim: to estimate the probability of disease development Target population: general population Baseline data and study design: cohort study population Statistical method: Cox-proportional hazard model	Framingham heart score [2,3]
Competing RA model	Aim: to estimate the probability of disease development Target population: general population Baseline data and study design: RR estimates from a case-control study; general population-based incidence and mortality Statistical method: Logistic regression model for computing RR estimates and complex integral calculus equations applied to age-specific data	BARAT [4] Colon cancer [5,6] Lung cancer [7,8] Prostate cancer [9] Melanoma [10]
Relative RA model	Aim: to estimate the probability of disease development Target population: general population Baseline data: risk point determined by experts' consensus based on previous representative studies; population based average risk (SEER) Statistical method: the score summed individual risk factor risk estimates divided by the population average projects to individual 10-year cumulative risk by multiplying population average cumulative risk	Harvard Cancer Risk Index [11]
RA model for heredity	Aim: to estimate the probability of disease development Target population: specific population with heredity or family history of a certain disease Baseline data and study design: family history in the specific populations based on pedigrees and/or prevalence of gene mutation Statistical method: Penetrance estimation and disease developing risk calculation according to gene mutation	With genetic information: BRCAPRO [12] PancPRO [13] Without genetic information: Claus model [14] IBIS (Tyrer-Cuzick) model [15]
RA model using probability that a given individual has the disease	Aim: to estimate the conditional probability given a certain disease Target population: general population or specific population exposed to specific factors such as toxicants Baseline data and study design: case-control (diagnosis, high risk group detection) or case series data (prognosis) Statistical method: logistic regression model	Clinicogenomic model for lung cancer diagnosis [16] Model for silica exposed population [17] Model for occupational asthma [18]

probability that is calculated from the cohort population is projected onto the general population. An example is the Framingham and National Cholesterol Education Program tools [2] developed from the Framingham risk score in 1948 [3] to predict a person chance of having a heart attack [7].

The second model uses a competing risk analysis. It is based on data from a case-control study design for relative risk (RR) estimates and general population statistics as incidence and mortality rates. Disease probability in the future is calculated using integral calculus equations based on age-specific data such as joint odds ratios calculated by a combination of risk factors in case-control studies and the incidence rates of a certain disease along with the mortality rates of other diseases. The target population is the general population. Many RA models use the competing risk method. A well-known example is The Breast Cancer Risk Assessment Tool (BCRAT) designed by the National Cancer Institute and the National Surgical Adjuvant Breast and Bowel Project that aimed to estimate a woman's risk of developing invasive breast cancer from 1989 [4]. The initial model was intended for Caucasian women but has been updated for African-American women using the Contraceptive and Reproductive Experiences study [20].

The third type is the relative risk model of the 'Harvard Cancer Risk Index' reported in 2000 [11]. The Risk Index Working Group at Harvard University used group consensus among researchers to classify risk factors into definite, probable, and possible causes of cancer. Baseline information was collected from representative studies and a risk score was developed through researchers' consensus. The relative risks due to individual risk factors from the representative data were converted to cancer risk points according to the strength of the causal association. The risk estimates were then totaled. The total was divided by the population average risk of cancer for a person with the same age and sex and multiplied by the average 10-year risk for disease from the Surveillance, Epidemiology and End Results (SEER) data that calculated the 10-year cumulative risk in the US population [11].

The fourth model is targeted to specific populations such as specific heredity factors or families. An example is the BACAPRO model for breast cancer risk prediction due to BRCA 1/2 gene mutations and a family history of breast and ovarian cancer [12].

Although several risk prediction models exist at present, accurately predicting an individual's risk for disease remains a challenge. Several papers have reported that traditional RA tools do not accurately predict coronary heart disease [21]. Substantial limitations have also been reported when these tools are used to guide individual therapy [22,23]. There are significant limitations in breast cancer models, such as their ability to predict breast cancer risk accurately when allowing for individualization of management strategies as well as model limits in ensuring that appropriate and autonomous decision-making takes place [24].

Advances in information pertaining to genetic susceptibility factors are expected to be most responsible for the differences in individual disease risk. It is anticipated that revised RA risk models which incorporate gene information and molecular markers will improve prediction accuracy. Dr. Mitchell H. Gail at the US National Cancer Institute added seven single-nucleotide polymorphisms (SNPs) that had previously been associated with breast cancer to the BCRAT. Although the model that included the SNPs provided less discriminatory accuracy than the BCRAT, it showed the potential to improve the discriminatory accuracy of the BCRAT modestly (the area under the receiver operating characteristic curve from 0.607 to 0.632) [25]. Dr. Gail showed small improvements in the

benefit when deciding whether to take tamoxifen for prevention and whether to recommend mammogram screening in the BCRAT plus 7 SNPs model compared with the BCRAT [26]. New RA models with genetic information will effectively characterize substantial high-risk individuals in populations exposed to silica [17] and occupational asthmatic factors [18]. Moreover, experience in molecular genetic testing applied to patients' prognosis and diagnosis provides knowledge that molecular genomic markers play an important role in the pre-detection and prognosis of individual diseases in a variety of clinical contexts, which heralds the future of genomics-based Personalized Predictive Preventive Medicine as idealized by Dr. Francis Collins. Two examples are as follows. MammaPrint is a molecular genomic diagnostic test that assesses the risk of breast cancer progression and drug response. It is more accurate in predicting a patient's prognosis relative to clinicopathological factors such as the tumor size and lymphatic invasion [27]. The second example is a combined clinic-genomic model that adds gene expression markers to traditional clinical factors for lung cancer diagnosis. It can diagnose lung cancer with higher validity (nearly 100% sensitivity and 91% specificity), even when the physician's diagnosis is uncertain [16].

We expect improvements in the power of elucidation on human risk assessment by adding molecular genomic factors. These new RA models, which include molecular genomic markers, will increase predictability by reducing misclassification bias through more valid exposure biomarkers.

PHARMACOGENOMICS IN MOLECULAR AND GENOMIC EPIDEMIOLOGY

Pharmacogenetics or pharmacogenomics (PGx) is the study of variations in DNA and RNA related to drug response, including drug absorption, disposition, and effect (efficacy and safety) [28]. Given that researchers reported an inherited basis for drug-response phenotypes in enzyme activity studies in the 1950s (reviewed in [29]), pharmacogenetic biomarkers have the potential to identify responders and non-responders, avoid toxicity and adjust the dosage of drugs to optimize efficacy and safety for each patient by providing an integrated approach to developments in genomics and associated technological innovations. This has implications not only for individual treatment but for drug discovery and development methods [30].

Early PGx studies were hypothesis-based candidate gene approaches that resulted in a few striking examples, such as 6-mercapto purine and *TPMT* [31], clopidogrel and *CYP2C19* [32], irinotecan and *UGT1A1* [33] and carbamazepine and the *HLA-B*1502* allele [34]. The simple methodology and statistical analysis used in conjunction with a comparably small sample size is the advantage of the candidate approach; however, it is rare to find strong and monogenic polymorphism candidates for drug response, low power for polygenic traits, and limited knowledge of gene and polymorphic function and drug pathways [35], thereby missing the true candidate gene and variation. As an extension of the candidate gene approach, a pathway-based genotype approach evaluates the combined effects of multiple genes in the same and/or different pathways to identify more sensitive and specific predictor profiles of drug response, such as *CYP2C9* as a metabolizing enzyme and *VKORC1* as a drug target for warfarin dosing [36], and DNA repair genes (*XPD*, *ERCC1*, and *XRCC1*) for clinical outcome of cisplatin-based chemotherapy [37] to amplify the modest effects of individual polymorphisms and enhance the predictive power. However, this remains limited by previous knowledge in its selection of the genes and pathways for drugs. It is often difficult to analyze the gene-gene interaction based on the *a priori* biological pathway due to the uncertainty to assume equal weight for each allele and the arbitrary assignment of the unfavorable allele in a case in which no information pertaining to functional inferiority [38]. In contrast, the genome-wide scanning approach is a hypothesis-generating design that does not depend on current knowledge to conduct a non-biased global genome assessment. Although the modest effect size of common variants, even in combination, leads to an argument of biological plausibility [39,40], the GWAS approach expects to discover the biologic pathways in polygenic traits [41]. Through the progress of human genetic variation and the haplotype map (known as the HapMap) coupled with rapid improvements in genotyping technology and analysis, 416 genome-wide association studies (GWAS) attempted to assay at least 100,000 SNPs in the first stage, and SNP-trait associations with p-values of less than 1.0×10^{-5} have been published thus far, reporting 400 novel GWAS loci in 75 diseases in a catalog published by the Human Genome Research Institute [42]. However, it appears premature to launch a GWAS in PGx because drug response as an outcome is more complex and the sample size with specific types of responder/non-responder or adverse reactions is much smaller compared to a typical disease-association study. Therefore, tremendous international collaboration is required in terms of a treatment regimen and drug response, such as that for *SLCO1B1* and statin-induced myopathy, as conducted in two sets of patient and control groups from large trials of approximately 12,000 and 20,000 participants from the SEARCH Collaborative Group [43].

In the guidance for industry pharmacogenomic data submissions issued by the US FDA, a valid genomic biomarker is defined as a biomarker that is measured in an analytical test system with well-established performance characteristics and for which there is an established scientific framework or body of evidence that elucidates the physiologic, toxicologic, pharmacologic, or clinical

Table 2. Valid pharmacogenetic biomarkers among drug labels approved by the FDA

Valid biomarker	Drug	Drug label information / possible mechanism
CYP2C19	Clopidogrel	CYP2C19 poor metabolizer is associated with lower systemic exposure to the active metabolite
CYP2C19	Voriconazole	*CYP2C19* variants (poor metabolizer or ultra-rapid metabolizer) with genetic defect leads to change in voriconazole exposure
CYP2C9	Celecoxib	CYP2C9 poor metabolizer may have abnormally high plasma levels due to reduced metabolic clearance of celecoxib
CYP2C9 and *VKORC1*	Warfarin	*CYP2C9*2* or *CYP2C9*3* alleles is associated with reduced warfarin clearance; *VKOR*C1 variant haplotypes in regulatory regions leading to variable expression with higher anticoagulation
CYP2D6	Atomoxetine	CYP2D6 poor metabolizer is associated with higher systemic exposure to atomoxetine
CYP2D6	Fluoxetine	Fluoxetine inhibits the activity of CYP2D6, and thus may make normal metabolizers resemble poor metabolizers
CYP2D6	Codeine	Ultra-rapid metabolizers is associated with higher systemic exposure to morphine, the active metabolite
DPD	Capecitabine	DPD deficiency increased side effect risk
Familial Hypercholestremia (deficiency, and/or mutation, of receptors for low density lipoprotein [LDL])	Atorvastatin	Altered HMG-CoA reductase activity; dosage adjustment for homozygous and heterozygous familial hypercholestremia
*HLA-B*1502* allele	Carbamazepine	Altered immunologic response leading to serious dermatologic adverse effect
*HLA-B*5701* allele	Abacavir	Altered immunologic response leading to hypersensitivity, lactic acidosis and severe hepatomegaly
KRAS mutation	Panitumumab	Lack of efficacy in colorectal cancer patients with *KRAS* mutations
NAT variants	Rifampin Isoniazid Pyrazinamide	Slow acetylation may lead to higher blood levels of the drug
TPMT	Azathioprine	Thiopurine methyltransferase deficiency or lower activity due to mutation increased exposure to drug
UGT1A1	Irinotecan	Nilotinib *UGT1A1* mutation with decreased expression is associated with higher systemic exposure to drugs

Adopted and modified from [47].

Table 3. Comparison of epidemiological study designs in pharmacogenetic studies

Observational study (case-series study)	Ancillary study for biomarkers to clinical trial	Randomized controlled trial for biomarkers
General population Real world treatment	Selected and uniform treatment regimen More careful outcome assessment Cost and time efficiency: already established collaborative setting	Can establish causality (prospective study) Treatment assignment based on the biomarkers More careful outcome assessment
Heterogeneous treatment regimen Heterogeneous outcome assessment Cannot establish causality	Selected population Cannot establish causality	Selected population Often small sample size

Adopted and modified from [48] with copyright permission.

significance of the test results [44]. There have been a limited number of potentially useful examples of pharmacogenomic DNA and RNA variant biomarkers approved by the FDA, although they are categorized as 'test recommended' or 'information only' (Table 2).

To overcome the challenges related to the applicability of genomic information as a valid biomarker in clinical practice, the study design needs to be considered. Most studies published were based on observational retrospective assessments of patients receiving treatment while the heterogeneity of the disease and treatment increased the likelihood of spurious findings. Although in the context of experimental designs (clinical trials), valid retrospective studies of a genomic biomarker require data from well-conducted randomized clinical trials; sample availability on a majority of patients to avoid selection bias; a prospectively stated hypothesis, analysis techniques, predefined and standardized assay and scoring system; and large sample size and power justification (reviewed in [45]). The gold standard is the prospective clinical trials that provide the setting for the evaluation of genomic biomarker validation, although this method is in general more expensive than an observational study. For example, the Trial Assigning Individualized Options for Treatment (TAILORx) was designed to integrate the 21-gene assay into the clinical decision-making process based on genomic biomarker recurrence scores in tamoxifen-treated patients with breast cancer [46]. Table 3 summarizes the general pros and cons of various study approaches.

Other issues as we move toward clinical integration of PGx information include the following: 1) defining drug response, suitably defined clinical endpoints (i.e., death), biomarkers or surrogate endpoints; 2) careful and sophisticated analysis of inter-individual variability in the demographic, environmental, lifestyle, and/or physiologic factors that affect drug response; 3) a cut off value for biomarkers based on supportive and rational analytical data and on the study design; 4) ethnic differences and population stratification; 5) the use of comprehensive inherited information such as epigenetic factors and copy number variation as well as genetic factors to improve sensitivity.

SUMMARY

Risk assessment and personalized medicine using molecular and genomic markers is a young field that holds considerable promise for contributions to healthcare by prevention,

earlier diagnosis, early intervention, effective interventions, and better outcomes. The major challenges associated with genetic biomarkers appear to be a lack of reproducibility and validation. Collaborative effort from various experts such as epidemiologists, clinicians, biologists, and biostatisticians is needed for further development and for the creation of novel approaches.

REFERENCES

1. Schulte PA, Waters M. Using molecular epidemiology in assessing exposure for risk assessment. *Ann N Y Acad Sci* 1999; 895: 101-111.
2. Wilson PW, D' Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998; 97(18): 1837-1847.
3. Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: The Framingham Study. *Am J Cardiol* 1976; 38(1): 46-51.
4. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989; 81(24): 1879-1886.
5. Selvachandran SN, Hodder RJ, Ballal MS, Jones P, Cade D. Prediction of colorectal cancer by a patient consultation questionnaire and scoring system: A prospective study. *Lancet* 2002; 360(9329): 278-283.
6. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Using risk for advanced proximal colonic neoplasia to tailor endoscopic screening for colorectal cancer. *Ann Intern Med* 2003; 139(12): 959-965.
7. Spitz MR, Hong WK, Amos CI, Wu X, Schabath MB, Dong Q, et al. A risk model for prediction of lung cancer. *J Natl Cancer Inst* 2007; 99(9): 715-726.
8. Bach PB, Kattan MW, Thornquist MD, Kris MG, Tate RC, Barnett MJ, et al. Variations in lung cancer risk among smokers. *J Natl Cancer Inst* 2003; 95(6): 470-478.
9. Fears TR, Guerry D 4th, Pfeiffer RM, Sagebiel RW, Elder DE, Halpern A, et al. Identifying individuals at high risk of melanoma: A practical predictor of absolute risk. *J Clin Oncol* 2006; 24(22): 3590-3596.
10. Optenberg SA, Clark JY, Brawer MK, Thompson IM, Stein CR, Friedrichs P. Development of a decision-making tool to predict risk of prostate cancer: The Cancer of the Prostate Risk Index (CAPRI) test. *Urology* 1997; 50(5): 665-672.
11. Colditz GA, Atwood KA, Emmons K, Monson RR, Willett WC, Trichopoulos D, et al. Harvard report on cancer prevention volume 4. Harvard Cancer Risk Index: Risk Index Working Group, Harvard Center for Cancer Prevention. *Cancer Causes Control* 2000; 11(6): 477-488.
12. Berry DA, Parmigiani G, Sanchez J, Schildkraut J, Winer E. Probability of carrying a mutation of breast-ovarian cancer gene BRCA1 based on family history. *J Natl Cancer Inst* 1997; 89(3): 227-238.
13. Wang W, Chen S, Brune KA, Hruban RH, Parmigiani G, Klein AP. PancPRO: Risk assessment for individuals with a family history of pancreatic cancer. *J Clin Oncol* 2007; 25(11): 1417-1422.
14. Claus EB, Risch N, Thompson WD. Autosomal dominant inheritance of early-onset breast cancer: Implications for risk prediction. *Cancer* 1994; 73(3): 643-651.
15. Cuzick J, Forbes J, Edwards R, Baum M, Cawthorn S, Coates A, et al. First results from the International Breast Cancer Intervention Study (IBIS-I): A randomised prevention trial. *Lancet* 2002; 360(9336): 817-824.
16. Beane J, Sebastiani P, Whitfield TH, Steiling K, Dumas YM, Lenburg ME, et al. A prediction model for lung cancer diagnosis that integrates genomic and clinical features. *Cancer Prev Res (Phila Pa)* 2008; 1(1): 56-64.
17. McCanlies E, Landsittel DP, Yucesoy B, Vallyathan V, Luster ML, Sharp DS. Significance of genetic information in risk assessment and individual classification using silicosis as a case model. *Ann Occup Hyg* 2002; 46(4): 375-381.
18. Demchuk E, Yucesoy B, Johnson VJ, Andrew M, Weston A, Germolec DR, et al. A statistical model for assessing genetic susceptibility as a risk factor in multifactorial diseases: Lessons from occupational asthma. *Environ Health Perspect* 2007; 115(2): 231-234.
19. National Heart Lung and Blood Institute. Risk Assessment Tool for Estimating Your 10-year Risk of Having a Heart Attack. Bethesda, MD: National Heart Lung and Blood Institute; 2004 [cited 2009 Oct 10]. Available from: URL:http://hp2010.nhlbihin.net/ATPiii/calculator.asp.
20. National Cancer Institute. Breast cancer risk assessment tool. Bethesda, MD: National Cancer Institute; 2007 [cited 2009 Oct 5]. Available from: URL:http://www.cancer.gov/bcrisktool/.
21. Johnson KM, Dowe DA, Brink JA. Traditional clinical risk assessment tools do not accurately predict coronary atherosclerotic plaque burden: A CT angiography study. *AJR Am J Roentgenol* 2009; 192(1): 235-243.
22. Akosah KO, Schaper A, Cogbill C, Schoenfeld P. Preventing myocardial infarction in the young adult in the first place: How do the National Cholesterol Education Panel III guidelines perform? *J Am Coll Cardiol* 2003; 41(9): 1475-1479.
23. Nasir K, Michos ED, Blumenthal RS, Raggi P. Detection of high-risk young adults and women by coronary calcium and National Cholesterol Education Program Panel III guidelines. *J Am Coll Cardiol* 2005; 46(10): 1931-1936.
24. Bellcross CA, Lemke AA, Pape LS, Tess AL, Meisner LT. Evaluation of a breast/ovarian cancer genetics referral screening tool in a mammography population. *Genet Med* 2009. Epub 2009 Sep 11.
25. Gail MH. Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. *J Natl Cancer Inst* 2008; 100(14): 1037-1041.
26. Gail MH. Value of adding single-nucleotide polymorphism genotypes to a breast cancer risk model. *J Natl Cancer Inst* 2009; 101(13): 959-963.
27. Buyse M, Loi S, van' t Veer L, Viale G, Delorenzi M, Glas AM, 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(17): 1183-1192.
28. Food and Drug Administration, HHS. International Conference on Harmonisation; Guidance on E15 pharmacogenomics Definitions and Sample Coding; Availability. notice. *Fed Regist* 2008; 73(68): 19074-19076.
29. Meyer UA. Pharmacogenetic: five decades of therapeutic lessons from genetic diversity. *Nat Rev Genet* 2004; 5(9): 669-676.
30. Roses AD. Pharmacogenetics in drug discovery and development: A translational perspective. *Nat Rev Drug Discov* 2008; 7(10): 807-817.
31. Relling MV, Hancock ML, Rivera GK, Sandlund JT, Ribeiro RC, Krynetski EY, et al. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. *J Natl Cancer Inst* 1999; 91(23): 2001-2008.
32. Simon T, Verstuyft C, Mary-Krause M, Quteineh L, Drouet E, Meneveau N, et al. Genetic determinants of response to clopidogrel and cardiovascular events. *N Engl*

J Med 2009; 360(4): 363-375.
33. Innocenti F, Kroetz DL, Schuetz E, Dolan ME, Ramirez J, Relling M, et al. Comprehensive pharmacogenetic analysis of irinotecan neutropenia and pharmacokinetics. *J Clin Oncol* 2009; 27(16): 2604-2614.
34. Chung WH, Hung SI, Hong HS, Hsih MS, Yang LC, Ho HC, et al. Medical genetics: A marker for Stevens-Johnson syndrome. *Nature* 2004; 428(6982): 486.
35. Wu X, Gu J, Spitz MR. Strategies to identify pharmacogenomic biomarkers: Candidate gene, pathway-based, and genome-wide approaches. In: Innocenti F, editor. *Genomics and Pharmacogenomics in Anticancer Drug Development and Clinical Response*. New Jersey: Humana Press; 2009. p. 353-370.
36. Klein TE, Altman RB, Eriksson N, Gage BF, Kimmel SE, Lee MT, et al. Estimation of the warfarin dose with clinical and pharmacogenetic data. *N Engl J Med* 2009; 360(8): 753-764.
37. Quintela-Fandino M, Hitt R, Medina PP, Gamarra S, Manso L, Cortes-Funes H, et al. DNA-repair gene polymorphisms predict favorable clinical outcome among patients with advanced squamous cell carcinoma of the head and neck treated with cisplatin-based induction chemotherapy. *J Clin Oncol* 2006; 24(26): 4333-4339.
38. Wu X, GU J. Pharmacogenetics in cancer chemotherapy. In: Rebbeck TR, Ambrosone CB, Shields PG, editors. *Molecular Epidemiology: Applications in Cancer and Other Human Diseases*. New York: Informa Heathcare; 2008. p. 113-128.
39. Wang WY, Barratt BJ, Clayton DG, Todd JA. Genome-wide association studies: Theoretical and practical concerns. *Nat Rev Genet* 2005; 6(2): 109-118.
40. Ziegler A, Konig IR, Thompson JR. Biostatistical aspects of genome-wide association studies. *Biom J* 2008; 50(1): 8-28.
41. Hirschhorn JN. Genomewide association studies: Illuminating biologic pathways. *N Engl J Med* 2009; 360(17): 1699-1701.
42. National Human Genome Research Institute. A Catalog of Published Genome-Wide Association Studies. Bethesda, MD: National Human Genome Research Institute; 2009 [cited 2009 Oct 5]. Available from: URL: http://www.genome.gov/26525384.
43. Link E, Parish S, Armitage J, Bowman L, Heath S, Matsuda F, et al. SLCO1B1 variants and statin-induced myopathy: A genomewide study. *N Engl J Med* 2008; 359(8): 789-799.
44. Center for Drug Evaluation Research, Center for Biologic Evaluation Research, Center for Devices and Radiological Health. Guidance for Industry Pharmacogenomic Data Submissions. Silver Spring, MD: Food and Drug Administration; 2005 [cited 2009 Oct 5]. Available from: URL: http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm126957.pdf.
45. Mandrekar SJ, Sargent DJ. Clinical trial designs for predictive biomarker validation: Theoretical considerations and practical challenges. *J Clin Oncol* 2009; 27(24): 4027-4034.
46. Sparano JA, Paik S. Development of the 21-gene assay and its application in clinical practice and clinical trials. *J Clin Oncol* 2008; 26(5): 721-728.
47. Food and Drug Administration. Table of Valid Genomic Biomarkers in the Context of Approved Drug Labels. Silver Spring, MD: Food and Drug Administration; 2009 [cited 2009 Oct 5]. Available from: URL:http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm.
48. Ulrich CM, Ambrosone CB. Molecular epidemiology designs for prognosis. In: Rebbeck TR, Ambrosone CB, Shields PG, editors. *Molecular Epidemiology: Applications in Cancer and Other Human Diseases*. New York: Informa Heathcare; 2008: p. 41-52.