Clinical utility of polymorphisms in one-carbon metabolism for breast cancer risk prediction

Shaik Mohammad Naushad¹, Shree Divyya¹, Addepalli Pavani¹, Yedluri Rupasree¹, Raghunadha Rao Digumarti², Suryanarayana Raju Gottumukkala³, Vijay Kumar Kutala¹

¹Department of Clinical Pharmacology & Therapeutics, ²Department of Medical Oncology, ³Department of Surgical Oncology, Nizam’s Institute of Medical Sciences, Panjagutta, Hyderabad 500 082, Andhra Pradesh, India.

This study addresses the issues in translating the laboratory derived data obtained during discovery phase of research to a clinical setting using a breast cancer model. Laboratory-based risk assessment indicated that a family history of breast cancer, reduced folate carrier 1 (RFC1) G80A, thymidylate synthase (TYMS) 5’-UTR 28bp tandem repeat, methylene tetrahydrofolate reductase (MTHFR) C677T and catecholamine-O-methyl transferase (COMT) genetic polymorphisms in one-carbon metabolic pathway increase the risk for breast cancer. Glutamate carboxypeptidase II (GCPII) C1561T and cytosolic serine hydroxymethyl transferase (cSHMT) C1420T polymorphisms were found to decrease breast cancer risk. In order to test the clinical validity of this information in the risk prediction of breast cancer, data was stratified based on number of protective alleles into four categories and in each category sensitivity and 1-specificity values were obtained based on the distribution of number of risk alleles in cases and controls. Receiver operating characteristic (ROC) curves were plotted and the area under ROC curve (C) was used as a measure of discriminatory ability between cases and controls. In subjects without any protective allele, aberrations in one-carbon metabolism showed perfect prediction (C=0.93) while the predictability was lost in subjects with one protective allele (C=0.60). However, predictability increased steadily with increasing number of protective alleles (C=0.63 for 2 protective alleles and C=0.71 for 3 protective alleles). The cut-off point for discrimination was >4 alleles in all predictable combinations.

Models of this kind can serve as valuable tools in translational research, especially in identifying high-risk individuals and reducing the disease risk either by life style modification or by medical intervention.

Key words: breast cancer, one-carbon metabolism, polymorphism, odds ratios, receiver operating characteristic (ROC) curve, risk prediction

© 2011 Deccan College of Medical Sciences. All rights reserved.
In recent years, breast cancer has replaced cervical cancer as the most common cancer among Indian women. The incidence of breast cancer currently stands at 24.4 to 29.2 per 1,00,000 women. Breast cancer is a manifestation of abnormal genetic as well as epigenetic changes. Highly penetrant mutations account for a small proportion of breast cancer cases. Several studies were conducted globally, to explore the etiological factors contributing to breast cancer. With the elucidation of human genome sequence low penetrant genetic polymorphisms have become important predictive as well as prognostic markers for multi-factorial disorders.

One-carbon metabolism can influence both genetic and epigenetic changes as it plays a crucial role in DNA synthesis and DNA methylation.

One-carbon metabolism is characterized by sequential transfer of one-carbon moieties from one substrate to another. Dietary folate in the form of folyl polyglutamates is hydrolyzed by glutamate carboxypeptidase II (GCPII, MIM: 600934) an intestinal enzyme, to monoglutamate form thus facilitating the intestinal absorption of folate. Folate reductase enzyme reduces the folate into dihydrofolate (DHF) and tetrahydrofolate (THF). Tetrahydrofolate in the plasma is transported into red blood cells through reduced folate carrier 1 (RFC1, MIM: 600424), where it accepts methylene moiety from serine to form 5,10-methylene THF in the presence of cytosolic serine hydroxymethyltransferase (cSHMT, MIM: 182144). 5,10-methylene THF is a substrate for two rate-limiting enzymes i.e. thymidylate synthase (TYMS, MIM: 188350) and methylene tetrahydrofolate reductase (MTHFR, MIM: 607093). The former catalyzes the conversion of dUMP to dTMP while the latter catalyzes the FAD-dependent reduction of 5,10-methylene THF to 5-methyl THF. 5-methyl THF donates a methyl group for remethylation of homocysteine to methionine in the presence of methionine synthase (MTR, MIM: 156570)-methionine synthase reductase (MTRR, MIM: 602568) holoenzyme complex. Methionine is the precursor for the synthesis of S-adenosyl methionine (SAM), a universal methyl group donor that donates methyl group to DNA, proteins and catecholamines. Catechol estrogens generated by cytochrome P450 enzymes are detoxified by O-methylation in the presence of catecholamine methyl transferase (COMT) enzyme (Fig 1).

Any perturbation in this pathway can induce i) uracil misincorporation in DNA causing DNA damage; ii) aberrant DNA methylation (focal hypermethylation and global hypomethylation) that triggers activation of proto-oncogenes and inactivation of tumor suppressor genes; and iii) toxicity due to catechol estrogens.

Recently we reported that low folate intake, RFC1 G80A (rs1051266), TYMS 5'-UTR 28bp tandem repeat and MTHFR C677T polymorphisms increase risk for breast cancer in Indian women while cSHMT C1420T polymorphism conferred protection against breast cancer.

These observations were consistent with other epidemiological studies indicating an inverse association between breast cancer risk and dietary intake of green vegetables, white vegetables, mushrooms and folate. Low folate intake or carrier status for BRCA1 or prolonged exposure to the estrogens prior to first full term pregnancy has been found to increase the risk for the breast cancer synergistically in the subjects with MTHFR 677 T-allele. MTR A2756G (rs1805087) polymorphism has been reported to reduce the risk for the breast cancer in one study whereas, in another study, no such association has been observed. MTRR A66G (rs1801394) polymorphism has been shown to alter the susceptibility to colorectal cancer, acute leukemia, lung cancer and squamous cell carcinoma, however, no such association has been observed with breast cancer. A limited number of studies have been reported on the RFC1 G80A (rs1051266) and cSHMT C1420T (rs1979277) polymorphisms with relevance to the breast cancer.

In the past decade, genetic association studies and genome-wide association studies (GWAS) have identified several genetic polymorphisms that are associated with many complex diseases and traits. However, a significant challenge remains in evaluating whether a genetic risk factor profile is clinically useful, either in a public health setting or for personalized medicine. In the current study, we validate the clinical utility of polymorphisms in one-carbon metabolism in predicting the breast cancer risk.

Materials and methods

Subjects

A case-control study was conducted at Nizam’s Institute of Medical Sciences, Hyderabad, during the period of January 2009 to January 2010. A consecutive series of female patients aged 18-80 years with newly diagnosed and histologically confirmed breast cancer was recruited as cases (n=342). Controls were healthy volunteers comprising of hospital staff and their relatives with no history of any cancer or benign breast disease. Subjects with any inflammatory or any other malignant disease were excluded. The rationale for choosing the controls was a matching with cases for age (± 5yrs), ethnicity, region and linguistic origin. Ultimately, 253 controls were enrolled in the study. Informed consent was obtained from all the subjects. This study was approved by the Institutional Ethical committee of Nizam’s Institute of Medical Sciences, Hyderabad (EC/NIMS/767/2007, dated 05.09.2008).

Demographic data collection and dietary assessment

All the subjects were recruited after personal interviews by trained interviewers using a standardized questionnaire which included demographic characteristics, life style (diet, exercise, smoking, alcohol intake, tobacco chewing), reproductive history (age of menarche, age at the time of first full term pregnancy, breast feeding history, number of live births, number of miscarriages, menopausal status, use of oral contraceptives, hormone replacement therapy) and family history (breast cancer or any other malignancies). Measurements of height and weight were taken at the time of recruitment from all the subjects to calculate body mass index (BMI). Clinical details were obtained from medical records of the patients. The blood samples were collected during the first visit to Oncology clinic.

Genetic analysis

Whole blood samples were collected in EDTA from all the subjects and the buffy coat was used for the genomic DNA isolation by the method described by Salazar et al. The extracted DNA was analyzed for eight genetic polymorphisms using the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) and PCR-amplified fragment length polymorphism (PCR-AFLP) methods. The protocols for genetic analysis have been described elsewhere. Cases and controls were analyzed in the same set of PCR by the analyst blinded to case-control status. For all the genetic analyses, each PCR set was accompanied by a negative control without genomic DNA in order to check contamination of components. Activities of all the restriction enzymes were checked by digesting the plasmids with known restriction sites. For RFLP analysis, a positive control was included in each set, which ensures complete digestion. Genotypes in few specimens were randomly rechecked.
to rule out genotyping errors and 100% concordance was observed.

Statistical analysis

Fisher’s exact test was used for obtaining odds ratio (OR) and 95% confidence interval (CI) for each variable. Simple logistic regression analysis was used to ascertain whether there is any trend of increase or decrease in the risk with increase in the number of risk/protective alleles. The data on genetic polymorphisms was computed in 0, 1 and 2 format depending on number of variant alleles. Since GCPII C1561T and cSHMT C1420T polymorphisms were protective against breast cancer, their cumulative variant allele frequency was used to stratify the risk allele data. Family history of breast cancer was computed as 0 and 1 based on absence or presence of familial breast cancer. Cumulative sum of family history of breast cancer and high risk alleles in RFC1, TYMS 5’UTR, MTHFR and COMT loci was used as predictor variable. Presence or absence of the disease was used as outcome variable. For each predictor variable, sensitivity and specificity were calculated using Fisher’s exact test. Receiver operating Characteristic curve (ROC) was plotted using 1-specificity on X-axis and sensitivity on Y-axis for all the predictor variables. A profile with area under the curve, (C) = 0.5 showing no discriminatory ability whereas (C) = 1 has perfect discriminatory ability. Further, a plot of predictor variable with true positive and true negative rate was plotted to obtain cut-off point of predictor variable that has greater discriminatory ability.

Results

The distribution of protective and risk alleles in cases and controls was tabulated. Table I shows the consistent increase in risk with increase in number of variant alleles. The genotyping data was stratified into 0, 1, 2 and 3 based on number of protective alleles and for each stratum odds ratios and confidence intervals were plotted using number of risk alleles as predictors (Fig 2).

In subjects without protective alleles, > 4 cut-off of risk alleles clearly discriminated the cases and controls, however with limited sensitivity (0.13) and high specificity (1.00). The C value being 0.63 indicates mild discriminatory ability.

In subjects with three protective alleles, no one of the risk alleles showed statistically significant association. However, > 4 alleles cut-off showed moderate sensitivity (0.22) and high specificity (1.00). The C value being 0.71 indicates moderate discriminatory ability.

Discussion

The current study highlights the importance of communicating the disease risk based on genetic profiles beyond odds ratios in view of earlier studies that showed strong association for disease, but were not indicative predictive value. Recently, we reported the role of strong epistatic interactions between RFC1 G80A/ TYMS 5’-UTR 2R3R/ MTHFR C677T in altering the susceptibility to breast cancer and further demonstrated the protective role of cSHMT C1420T polymorphism. These observations are consistent with the functionality of these polymorphisms and their relevance in the pathogenesis of breast cancer. RFC1 G80A polymorphism impairs folate transport across the RBC membrane and in vitro studies indicated down-regulation of RFC1 under conditions of severe folate deprivation. TYMS and MTHFR enzymes compete for common substrate i.e. 5,10-methylene...
Fig 3. Receiver operating characteristic (ROC) curves for breast cancer prediction model. A, B, C and D represent ROC curves (sensitivity vs. 1-specificity) for high-risk alleles when the protective alleles are 0, 1, 2 and 3. The area under the curve (C) was used to discriminate the cases and controls. The C=0.93 for A curve showed perfect discrimination, C=0.60 for B curve showed poor predictability, C=0.63 for C curve showed mild predictability while C=0.71 for D curve showed moderate predictability.

THF for the synthesis of thymidylate and 5-methyl THF. TYMS 2R allele reduces TYMS expression 2-4 times while MTHFR C677T polymorphism results in a thermolabile MTHFR enzyme that loses active dimer form and gets dissociated into inactive monomers with subsequent loss in FAD-binding capacity. Decreased TYMS and MTHFR activities result in uracil misincorporation in DNA and impaired methylation thus affecting genome integrity and gene expression profiling; two hallmark changes that trigger carcinogenesis. The risk for breast cancer specifically is inflated further when there is COMT H108L polymorphism, which induces a labile COMT variant with a shorter half-life. Decreased catalytic activity of COMT and decreased availability of methyl group for O-methylation of catechol estrogens for the synthesis of methoxy estrogens result in accumulation of catechol estrogens. Catechol estrogens form highly reactive intermediates semiquinones and quinones that have a tendency to form adducts with DNA and superoxide ions generated through that process which oxidatively damage the DNA further. These mutagenic lesions trigger carcinogenesis further. cSHMT C1420T polymorphism was found to confer protection against breast cancer by increasing the plasma folate pool. This protection could be due to induction of futile folate cycle in which cSHMT catalyzes irreversible conversion of methylene THF to 5-formyl THF (futile folate cycle) in order to maintain the one-carbon homeostasis.

The major challenge remains in translating this laboratory derived observation to a clinical setting for the purpose of either using this information as predictive testing or for individual choice to change one’s lifestyle or to undertake a medical intervention to reduce disease risk. Individual odds ratios although useful in relative risk assessment, single locus investigations fail to ascertain the discrimination between those who will develop the disease and those who will not, as these polymorphisms are present in healthy individuals in a frequency >1%. Multi-locus investigations have advantage over single-locus investigations where genetic risk factor profile can be used to deduce the sensitivity vs. 1-specificity graphs for different combination of alleles so that area under the ROC curve can be used to validate the discriminatory ability of a con-
continuous biomarker. In the current study, we applied these statistics by incorporating a number of protective alleles as additional markers that distinguishes genetic risk factor profile in breast cancer.

The cumulative risk alleles were obtained by summing up the presence/absence of familial breast cancer, variant alleles in RFC1, TYMS 5'-UTR, MTHFR and COMT loci. GCP II C1561T and cSHMT C1420T polymorphisms being protective, their cumulative variant allele number is used to stratify the data.

In the absence of GCP II C1561T and cSHMT C1420T polymorphisms, the genetic risk factor profile perfectly discriminated the cases and controls and the cut-off point for discrimination was > 4 alleles. The sensitivity was 0.76 and specificity was 1.00. However in subjects with one protective allele, the discriminatory ability of genetic risk factor profile is lost (C=0.60). However, discriminatory ability increased steadily with increase in number of protective alleles (for 2 alleles, C=0.63; for 3 alleles, C=0.71). The cut-off point of discrimination in these subjects was >4 alleles. This information should prove very much useful in public health perspective and can be easily interpreted by physician based on ROC plots. These kinds of risk prediction models will be of potential benefit when there is availability of an intervention proven to reduce risk.

To conclude, the statistics used during the discovery phase of research such as odds ratios and p values are not the most appropriate measures for evaluating the predictive values of genetic risk profiles. Stratifying the data based on high-risk and low-risk alleles should be helpful in measuring sensitivity and specificity across different strata and the discriminatory ability of the test can be conveyed based on ROC curves.

Conflict of interest: None

Acknowledgments

This work was supported by the grant funded by Indian Council of Medical Research (ICMR), New Delhi (Ref No. 5/13/32/2007) and Lady Tata Junior Research Fellowship awarded to Ms. Shreedi Diviya.

References


