

## ABSTRACT

Breast cancer is escalating in Sri Lanka in such a pace that the country may soon face a serious burden in terms of mortality and morbidity. This increasing incidence of breast cancer is not explained well only by its conventional risk factors.

Mammography screening aims to reduce the disease burden from breast cancer by early detection and treatment of occult malignancies in women over 40 years of age. However, absence of a routine screening programme hinders Sri Lankan women from early detection. Identifying the at-risk women for breast cancer by a low-cost simple risk prediction tool would be of paramount importance.

The objectives of the study were to identify the risk factors of breast cancer specific for Sri Lanka, to develop and validate a risk prediction model for identifying the at-risk women for breast cancer and to assess the prevalence of 'increased risk for breast cancer' among women in the district of Colombo.

The study consisted of three components. Component one was a case-control study, with newly diagnosed cases of breast cancer (n=210) selected from surgical, chemotherapy and radiotherapy units at the National Cancer Institute, Maharagama, and unmatched controls with no breast cancer (n=206) selected from the breast clinic in the same hospital. Cases and controls were recruited using a non-probability sampling method. Interviews were carried out by a trained interviewer at the patient's residence using an interviewer-administered-questionnaire. Quality of diet related to cancer was assessed using a food frequency questionnaire, validated before the study proper. Lifetime total physical activity levels was assessed in relation to occupational, household and sports/recreational activities using Lifetime Total Physical Activity Questionnaire (LTPAQ), assessed for its reliability before the study proper. Anthropometric measurements such as body weight, standing height and waist circumference were measured by standardized instruments. In a selected sample (61 cases and 51 controls) of women who had been residing at the same residence for > 10 years and had never been employed, indoor nitrogen dioxide levels were measured using passive samplers twice over two weeks period.

Bivariate analysis followed by logistic regression analysis were carried out to identify the independent risk factors for breast cancer adjusted for confounding effect in the logistic regression model. Of the socio-demographic factors, low income of < Rs. 20,000 (OR: 7.3; 95% CI: 1.76, 30.31) was significantly associated with breast cancer. Of hormonal and reproductive factors, total duration of breast feeding less than 24 months (OR: 2.62; 95% CI: 1.49, 4.61) and menarche before the age of 12 years (OR: 4.31; 95% CI: 1.59, 11.69) were the significant risk factors. Of lifestyle related factors, sub-optimal consumption of anti-oxidants (OR: 2.56; 95% CI: 1.45, 4.53) and fibre (OR: 7.4; 95% CI: 3.19, 17.27) and 'low' lifetime total physical activity (OR: 2.45; 95% CI: 1.09, 5.51) were found to be significant risk factors. Of genetic factors, having a first degree relative with breast cancer (OR: 6.25; 95% CI: 2.61, 14.97) and of the co-morbid factors, having diabetes mellitus (OR: 2.44; 95% CI: 1.08, 5.51) were found to be significant factors. In addition to that presenting with a breast lump was a significant factor for breast cancer (OR: 19.7; 95% CI: 9.1, 42.5).

In the second component of the study, a risk prediction model was developed based on the results of component one. Age, age at menarche, age at first live birth, total duration of breast feeding and having a first degree relative with breast cancer were identified as the risk predictors in the model. Each predictor was weighted by its OR to develop individual scores, and a summary risk score was calculated by adding the individual scores. The risk prediction model was applied in an independent sample comprising 44 cases and 62 unmatched controls recruited from the National Cancer Institute, Maharagama, to assess its validity as a screening tool, as well as to derive a cut-off value using ROC curves to distinguish at-risk women for breast cancer from no-risk women. At a cut-off value of 8.5, the risk prediction model had a sensitivity of 88.6% and a specificity of 79.0%. The positive and negative predictive values were 75.0% and 90.7% respectively.

In the third component of the study, the prevalence of 'increased risk for breast cancer' among women was determined in the district of Colombo using the validated risk prediction model. A sample of 1372 women who were 40 years or above was selected using a multistage, stratified cluster sampling. The adjusted prevalence of 'increased risk for breast cancer' among women was 33.3% (95% CI: 32.1%, 35.8%).

It is concluded that the risk factors specific for breast cancer among Sri Lankan women are mainly related to reproductive life. However, significant proportion of breast cancer was explained by lifestyle related factors and this adds evidence to the already existing knowledge. In view of delayed presentations and relatively high at-risk population in district of Colombo, the validated risk prediction model should be promoted as a simple, low cost tool for identifying and prioritising these women for routine mammography screening.

**Keywords: Breast cancer, risk factors, screening, risk assessment, prevalence**