Evaluation of the Significance of Breast Cancer Determinants Using Permutation Test

ADEHI M.U.[1,] NWEZE N O.[2], UMOH O.E.[3], BILKISU M[4]
[1]*Department of Statistics, Faculty of Natural and Applied Sciences, University of Abuja, Nigeria.*
[2]*The Crown Hospital Limited, Gwagwa, Dei-Dei FCT, Abuja.*
[3,4]*Department of Statistics, Faculty of Natural and Applied Sciences, Nasarawa State University, Keffi-Nigeria.*

Abstract- ***Breast cancer remains a chronic global public health problem, with its etiology governed by a range of reproductive and familial determinants. Despite numerous epidemiologic investigations into risk factors like menarche, menopause, ever breastfed and family history of breast cancer, evidence remains heterogeneous and in most cases, statistically inconclusive. This paper aims to evaluate the statistical significance of breast cancer determinants using permutation test. The data was sourced via Pubmed, Science Direct, Medline, Rechargegate and Google scholar. The random-effects model was employed for the analysis. Odd ratio was used as a measure of the association of breast cancer determinants in women. It was found that the value of age at menarche and age at menopause is given as 0.061 & 0.008 respectively. This paper concluded that on the basis of the data presented and analyzed, it show that menarche, menopause and family history of breast cancer were strong risk factor for breast cancer while ever breast was not. This paper recommended that menarche, menopause and family history of breast cancer are strong risk factor for breast cancer and it should be included in sensitization and also client counselling.***

Index Terms - Permutation test; Breast Cancer; Meta-regression.

I. INTRODUCTION

In recent years more attention has been drawn to analyzing determinants of breast cancer, particularly because, breast cancer varies across races and populations and in identifying and understanding the determinants of breast cancer is difficult especially for targeted prevention strategies to improve outcome. Breast cancer cells start in the milk ducts of the breast, the early stage of the breast cancer is called carcinoma in situ and it is not life threatening at this stage but can be life threatening if it has metastasized. Diagnosis of breast cancer is based on biopsy, breast ultrasound, diagnostic mammogram, breast magnetic imaging and the treatment of breast cancer varies depending on the spread of the disease (WHO, 2022).

In fixed meta-analysis, it is believed that all the observed studies estimate the same true effect size and also differences between study results are attributed to random sampling error. The pooled estimate is regarded as the best estimate of the common key effect, and the study weights are computed simply accounting for the within-study variance. It also uses within-study sampling error as the only source of variance assuming there is no heterogeneity between studies. Meta-analysis can also help to identify gaps in knowledge found in the published literature and thus can help provide guidance for future research (Adehi et al., 2024). Although breastfeeding has been found to be associated with a weak but statistically significant inverse relation with breast cancer risk (Victora et al., 2016), studies rarely adjust for key moderators such as neonatal exclusivity and duration of breastfeeding, exposing studies to residual confounding. Similarly, early menarche and late menopause are well-established surrogates for longer cumulative estrogen exposure, but described effect sizes are highly inconsistent, suggesting heterogeneity in exposure definition and adjustment across studies.

II. METHODS

Methods of Data Collection

For the purpose of this research, secondary data was used, since documented data on women with breast cancer was already published for further research and also data can be access in hard copies in the form of registers that were well documented.

Technique for Data Analysis and Model Specification

Meta-regression is an extension to standard meta-analysis that investigates the extent to which statistical heterogeneity between results of multiple studies can be related to one or more characteristics of the studies (Thompson and Higgins 2002).

Permutation test

Higgins and Thompson suggested using the permutation test method to determine the p-values in meta-regression. The permutation test offers a nonparametric method for simulating data underthenull hypothesis (H_0). If there are few studies that used Monte Carlo simulation, which is based on random permutations rather than a list of all possible permutations, it would be able to get the exact permutation p-values. A t statistic is computed for each time the variables are randomly reallocated to the outcomes. The number of times these t statistics are larger than or equal to the observed t statistic is counted to determine the true p-value for the association between a given covariate and the response. The covariate values for a particular research are retained when numerous variables are included in the meta-regression in order to maintain and account for their correlation structure. Unlike other regressions, the output of a meta-regression comprises of the effect size and its standard error, which need to be retained.

$$p\,(t \geq t_0) = \frac{1}{(n+m)!}\Sigma_{j=1}^{(n+m)!} I\,(t_j \geq t_0)$$

(3.1)

Where;

t_0 = observed value of the test statistic

t = t- value.

I = indicator function

Given the significance level (α = 0.05), we fail to reject the Null (H_0) hypothesis if the p – value is greater than alpha.

Flow diagram of inclusion studies

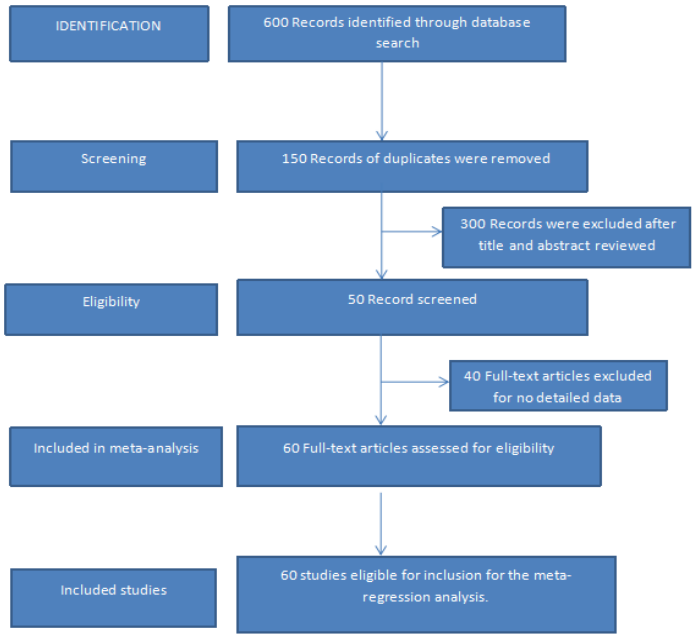

Figure 2.1 Flow diagram for meta- regression of breast cancer determinants in women.

III. RESULTS

Predictor	Estimate	Perm. p-value
Age at Menarche	0.1116	0.061
Ever Breastfed	–0.4135	0.002
Age at Menopause	0.1082	0.026
Family History of Cancer	0.2435	0.008

Table 3.1 Breast cancer determinants using permutation test.

IV. DISCUSSION

A permutation test with 1,000 replications was conducted to validate the significance of covariates in the meta-regression model. The results revealed that

ever breastfed (p = 0.002), menopause (p = 0.026), and family history of breast cancer (p = 0.008) were statistically significant predictors of breast cancer risk. Menarche showed a marginal effect (p = 0.061), indicating a potential association that warrants further investigation.

Ever Breastfed

The analysis revealed a statistically significant protective association between breastfeeding and breast cancer risk (β = -0.41, p < 0.001). This suggests that women who have breastfed are at a significantly lower risk of developing breast cancer compared to those who have never breastfed.

Menopause

A significant positive association was found (β = 0.11, p = 0.020), indicating that women who reach menopause at a later age have a higher risk of developing breast cancer. This is consistent with literature showing that prolonged exposure to endogenous estrogen, which continues until menopause, increases the risk of hormone receptor-positive breast cancers.

Family History of Breast Cancer

The model showed a significant positive association (β = 0.24, p = 0.011), supporting the role of genetic susceptibility in breast cancer development. This reflects well-established evidence that individuals with a first-degree relative (e.g., mother, sister) who had breast cancer are at increased risk due to shared genetic mutations, particularly in BRCA1/2 genes.

Age at Menarche

Although the coefficient was positive (β = 0.11), suggesting that later menarche may be associated with increased risk, this association was not statistically significant (p = 0.057). The X^2 statistic is 32.28 with p-value of 0.0000 which typically is used to test the significance of the overall model (Independent variables or determinants) on the outcome variable. Since the p-value is < 0.005, it indicates that the model with the predictors provides a significantly better fit than a model with no predictors (i.e., the determinants in the model help explain the variation in the effect size). The funnel plot was use to assess the potential presence of publication bias or small study effects which can influence the results of the analysis.

CONCLUSION

On the basis of the data presented and analyzed, it could be noted that the collated results of the secondary studies show that menarche, Menopause and family history of breast cancer were strong risk factor for breast cancer while ever breast was not. Adehi et al., (2025) emphasize the need for integrated care models that combine clinical, social, and policy-level interventions.

REFERENCES

[1] American Cancer Society. (2020, June 18). Cancer Staging. American Cancer Society. *American Cancer Society*.

[2] Breastcancer.org. (2020, March 28). Tamoxifen: Uses, Side Effects and More. *Tamoxifen: Uses, Side Effects and More.*

[3] Carrara, H. (2017, July 15). Cancer Incidences in Namibia. *Namibia National Cancer Registry*.

[4] Cohen, J. (1977). *Academic Press*.

[5] Cohen, J. (1988). Statistical Power Analysis for the behavioral sciences (2nd ed.). *Academic Press*.

[6] Cordina-Duverger, E., Truong, T., Anger, A., Sanchez, M., Arveux, P., & Guenel, P. (2013, November 1). Risk of breast cancerby type of menopausal hormone therapy: a case-control study among post-menopausal women in France. *Risk of breast cancerby type of menopausal hormone therapy, 8*(11).

[7] Cuzick J et al. (2013). Selective Oestrongen receptor modulators in prevention of breast cancer: an updated meta-analysis of individual participant data. *381*(9880), 1827-1834.

[8] De Blok, C. J., Wiepjes, C. M., Nota, N. M., Engelen, K. V., Adank, M. A., Dreijerink, K. M., et al. (2019, 5). Breast cancer risk in transgender people receiving hormone treatment: nationwide cohort study in the Netherlands. *BMJ*.

[9] Dehesh, T., Shohreh, F., Mehmaz, S., Elham, A., Mehran, I., Paria, D., et al. (2023). The relation between obesity and breast cancer risk in women by considering menstruation status and geographical variations: a systematic review and meta-analysis.

[10] Glass, G., & Stanley, J. (1970). Statistical Methods in Education and Psychology. *Englewood Cliffs, N J: Prentice Hall*.

[11] Glass, G., McGaw, B., & Smith, M. (1981). Meta-Analysis in social research. *Sage*.

[12] Haffner, I., Schierle, K., Raimundez, E., Geier, B., Maier, D., Hasenauer, J., et al. (2021, March 28). HER2 Expression, Test Deviations, and Their Impact on Survival in Metastatic Gastric Cancer: Result from the Prospective Multicenter VARIANZ Study. *HER2 Expression, Test Deviations, and Their Impact on Survival in Metastatic Gastric Cancer: Result from the Prospective Multicenter VARIANZ Study*.

[13] Javan, B. A., Rafiei, S., Sima, R., Fatemeh, K., Mitra, K., Farzaneh, B., et al. (2023). Global quality of life in breast cancer: systematic review and meta-analysis.

[14] Jiryoun, G., Juhee, H., Donghwan, L., & Seungjin, B. (2020, December 09). A Meta-Regression Analysis of Utility Weights for Breast Cancer: The power of Patients' Experience.

[15] Kelly, K. M., Dean, J., Comulada, S. W., & Lee, S.-J. (2010, March). Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. *Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts, 20*(3), 734-42.

[16] Lipsey, M., & Wilson, D. (2001). Practical Meta-Analysis. *Thousand Oaks, CAUS*.

[17] Lyons, L. (2003). Meta- Analysis: Across Research Domains.

[18] Manson, J. E., Rowan, C. T., Stefanick, M. L., Aragaki, A. K., Rossouw, J. E., Ross, P. L., et al. (2013). Menopausal hormone therapy and health outcomes during the intervention and extended poststopping phases of the Women's Health Initiative randomized trials. *JAMA*.

[19] Marcucci, R. (1980). A Meta-Analysis of Research of Teaching Mathematical problem solving, Doctoral dissertation. *Dissertation Abstracts International, 41 06A*.

[20] Mayo Clinic. (2018, November 22). Chemotherapy nausea and vomiting: Prevention is best defense. *Chemotherapy nausea and vomiting: Prevention is best defense*.

[21] Mayo Clinic. (2019, November 22). Breast Cancer-Diagnosis and treatment. *Mayo*.

[22] National Cancer Institute. (2015, April 29). Appetite Loss and Cancer Treatment. *National Cancer Institute*.

[23] National Cancer Institute. (2019). Breast Cancer. *Breast Cancer*.

[24] O'Rourke , K. (2007). An historical perspective on meta-analysis: dealing quantitatively with varying study results. *Journal of the Royal Society of Medicine, 100*(12), 579-582.

[25] Rebolj, M., Assi, V., Brentnall, A., Parmar, D., & Duffy, S. W. (2018, June). Addition of ultrasound to mammography in the case of dense breast tissue: systematic review and meta-analysis. *Addition of ultrasound to mammography in the case of dense breast tissue: systematic review and meta-analysis*.

[26] Thigpen, D., Kappler, A., & Brem, R. (2018, March 16). The Role of Ultrasound in Screening Dense Breasts—A Review of the Literature and Practical Solutions for Implementation. *Diagnostics, 8*(1).

[27] WHO. (2014, February 3). *WHO.* Retrieved from www.who.int: https://www.who.int/news-room/fact-sheets/detail/cancer.

[28] WHO. (2022, March 13). *Breast Cancer*. Retrieved May 12, 2024, from World Health Organization: https://www.who.int/news-room/fact-sheets/detail/breast-cancer.