Association between Breast Cancer and Hepatitis C Virus with Age Controlling

Amy Traianou
Department of Statistics
University of Connecticut

November 14, 2022

Abstract

In Egypt Hepatitus C virus infection is a common problem that often impedes treatment of Breast Cancer in female patients. While previous work has analyze an associaiton between HCV seropositivity and breast cancer diagnosis, this paper will focus on the potential confounding of age on the association. We will use Fisher's Exact Test to estimate the crude estimate of the association and the Mantel-Haenszel method to stratify the data between ages. The possible effect of age in the association can lead to improved care for patients.

KEYWORDS: Breslow-Day; Fisher's Exact Test; Mantel-Haenszel.

1 Introduction

This paper will explore the risk factors associated with breast cancer in women. Breast cancer is extremely prevalent and researchers are constantly trying to determine risk factors to identify women for preventative exams. There are many studies attempting to determine if there is an association between certain risk factors and breast cancer including chronic hepatitis C infection (Larrey, 2010) and hepatitis B (Adhikari et al., 2016).

Specifically, this paper will examine if there is an association between testing positive for hepatitis C virus and breast cancer while controlling for age in Egyptian populations. Both Hepatitis C and breast cancer are extremely prevalent in Egypt (Hussein et al., 2021). Additionally, research has found that women who are seropositive for HCV have more agressive forms of breast cancer. However, there is minimal research concerning the effect of age on the associaiton.

Recently on October 6th, Egypt and Qatar agreed to collaborate in the health sector and use each other's expertise. The health ministers specifically mentioned research in Hepatitis C and breast cancer as an area of interest (Online, 2022). This meeting emphasizes the importance of diving into Hepatitus C Virus and it's association with breast cancer. There have been conflicting foundings so studies should indenify additional variables that could contribute to better understanding the relationship. For that reason, this paper is focusing on controlling for age of female patients.

The rest of the paper is organized into several sections. Section 2 describes the study done in 2020 to test if HCV seropositivity is associated with breast cancer diagnosis in Egyptian population. Section 3 describes the following statistical methods; Fisher's Exact Test, Breslow-Day test, and Mantel-Haenszel method. Section 4 applies the methods to data from the study, including stratafied between ages. Section 5 concludes the paper.

2 Data

The data used in this paper is from a 2020 retrospective case-control study based on Egyptian female populations (University, 2020), sponsered by Mansoura University in Egypt.

The study group consists of 405 patients treated at the Oncology Center - Mansoura University in the past 10 years. In order to be included, invasive breast cancer must be biopsy-proven. Patients with unknown viral marker status, multiple cancer diagnoses and virus-unrelated hepatic pathology were all excluded.

The control group consists of data from a previous study, conducted from 2015-2017. There are 145 females from the same geographic distribution, all with no previous cancer diagnosis. In both groups, all patients are above the age of 18. For this paper, patients are stratified based on age- older patients are considered 45 years or above and younger patients are below 45 years. The data is summarized in the following 2x2 contingency tables.

2x2 Contingency Table for all patients:

	Breast Cancer	No Cancer Diagnosis	Total
HCV Seropositive	88	15	103
HCV Seronegative	317	130	447
Total	405	145	550

2x2 Contingency Table for pateints younger than 45:

	Breast Cancer	No Cancer Diagnosis	Total
HCV Seropositive	17	2	19
HCV Seronegative	110	91	201
Total	127	93	220

2x2 Contingency Table for patients older than 45:

	Breast Cancer	No Cancer Diagnosos	Total
HCV Seropositive	71	13	84
HCV Seronegative	207	39	246
Total	278	52	330

3 Methods

To test the association of breast cancer and Hepatitis C seropositivity, one can use the chi-square or Fisher's Exact test, depending on the conditions from the sample size (Warner, 2013). Fisher's exact test is useful for when the normality assumption is violated and the expected values of the 2x2 table are too small. The test uses the hypergeometric distribution to test if the probabilities are the same between the two groups. Thus, we can determine if there is more of a risk of breast cancer for those with seropositive Hep C. In the data section, the data was summarized into 2x2 contingency tables. The generalzied table is shown in Figure 1. The first step is calculate the probability of the original table occuring, using the hypergeometric pdf:

$$P_a = [\frac{(a+b+c+d)a!b!c!d!}{(a+c)!(b+d)!(a+b)!(c+d)!}]^{-}1 \tag{1}$$

Based on the alternative hypothesis, you then calculate the probability of all more extreme tables based on the a value. In this paper, p1 is the probability of having breast cancer given that the patient is seropositive and p2 is the probability of having breat cancer given that the patient is seronegative. Then, we want to determine if there is an association, specifically if the risk of having breast cancer is higher for those who are seropositive. The alternative hypothesis is that p1 is greater than p2- which will be tested in the application section.

After calculating the crude estimate, a Breslow-Day test will test for homogeneity of odds ratios between the age stratified data. The generalized form of stratified 2x2 contingency

	$Y=1$	$Y=0$	Total
$X=1$	A	C	A+C
$X=0$	B	D	B+D
Total	A+B	C+D	N

Figure 1: 2x2 Contingency Table

	$Y=1$	$Y=0$	Total
$X=1$	A_i	C_i	$A_i + C_i \equiv M_{1i}$
$X=0$	B_i	D_i	$B_i + D_i \equiv M_{0i}$
Total	$A_i + B_i \equiv R_{1i}$	$C_i + D_i \equiv R_{0i}$	N_i

Figure 2: 2x2 Contingency Table with Stratified Data

tables is shown in Figure 2. The Breslow-Day test statistic will be computed using SAS with the formula:

$$\chi^2_{BD} = \sum_{i=0}^{s} \frac{[A_i - E(A_i|\hat{OR}_{M-H})]^2}{Var(A_i|\hat{OR}_{M-H})} \tag{2}$$

Based on the results of the Breslow Day test for homogeneity of odds ratios, the Mantel-Haenszel test will compute a common odds ratio and test if there is an association between seropositivity and breast cancer while controlling for age. The Mantel-Haenszel test statistic is:

$$\chi^2_{M-H} = \frac{(|\sum_{i=0}^{s} A_i - \sum_{i=0}^{s} E(A_i)| - 0.5)^2}{\sum_{i=0}^{s} V(A_i)} \tag{3}$$

The expected value and variance of A for the ith strata is

$$E(A_i) = \frac{M_{i1}R_{1i}}{N_i}$$

$$V(A_i) = \frac{R_{1i}R_{0i}M_{1i}M_{0i}}{N_i^2(N_i - 1)}.$$

4 Application

Fisher's exact test for the pooled data of all patients tested

$$H_0 : p_1 = p_2, H_a : p_1 > p_2$$

at a 0.05 level of confidence.

The test resulted in a p-value of 0.0027, which is less than alpha = 0.05. Thus, the null hypothesis is rejected and there is evidence suggesting patients who are seropositive have a higher breast cancer rate than those who are seronegative.

Next, the Breslow Day test checks for homogeneity between the stratified ages.

$$H_0 : OR_1 = OR_2, H_a : OR_1 \neq OR_2$$

at a 0.01 level of confidence.

The test resulted in a p-value of 0.0133, which is greater than the alpha value. Thus, we can conclude the strata have similar odds ratios and proceed with the Mantel-Haenszel procedure.

The Mantel-Haenszel test produced a common odds ratio of 1.6847 and a 95 percent confidence interval of (0.9283, 3.0574). The interval contains 1 so after controlling for age, the association is not supported.

However, the individual odds ratios for age groups are worth noting. For the older group (45 and older) the odds ratio estimate and interval are 1.0290 and (0.5197, 2.0374). The odds ratio estimate and interval for the younger group (less than 45) are 7.0318 and (1.5828, 31.2399). Therefore, the association between seropositivity and breast cancer is much stronger for females less than 45 years.

5 Discussion

Based on the analysis, it is evident that for women under the age of 45, seropositivity and breast cancer are strongly associated. As being HCV positive can complicate treatment of cancer, this is critical information. In Egypt, breast cancer remains the most common form or cancer and high levels of hepatitus C continue. This study was retrospective so analysis was limited to observing odds ratios and the sample size was quite small. Thus, future studies should aim to be prospective cohort studies so researchers can observe how seropisitivity and breast cancer interact over time. Additionally, age is just one factor that can affect the associaiton between hepatitus C and breast cancer. Future studies should dive deeper into more factors such as weight, reproductive history, and alcohol consumption. Addiitonally, future studies can explore how seropositivity affects tumor size, disease agression, and more. As breast cancer is so prevalent in Egypt, further understanding the risk factors is becoming even more importnant.

References

Adhikari, V. P., L.-J. Lu, and L.-Q. Kong (2016). Does hepatitis b virus infection cause breast cancer? *Chinese Clinical Oncology 5*(6).

Hussein, O., E. M. El-Ghitany, M. Omran, G. Matariek, E. A. Elbadaly, R. Hamdy, A. Gamal, M. M. Zayed, A. Nasr, O. Hamdy, M. Elbasiony, and K. Abdelwahab (2021). High seroprevalence of hepatitis c virus antibody in breast cancer patients in egypt. *Breast Cancer: Basic and Clinical Research 15*, 117822342110024.

Larrey, D. (2010). Is chronic hepatitis c virus infection a risk factor for breast cancer? *World Journal of Gastroenterology 16*(29), 3687.

Online, A. (2022). Egypt, qatar agree on exchanging expertise in the health field.

University, M. (2020). Association of hepatitis c virus with breast cancer. https://www.clinicaltrials.gov/ct2/show/study/NCT04090164?recrs=e&rslt=With&type=Obsr&draw=13&rank=112.

Warner, P. (2013). Testing association with fisher's exact test. *BMJ Sexual & Reproductive Health 39*(4), 281–284.