

# Association between Breast Cancer and Hepatitis C Virus with Age Controlling

Amy Traianou  
Department of Statistics  
University of Connecticut

November 13, 2022

### **Abstract**

The Kolmogorov–Smirnov (KS) test is one of the most popular goodness-of-fit tests for comparing a sample with a hypothesized continuous distribution. Nevertheless, it has often been misused.

**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 (?) and hepatitis B (?).

Specifically, this paper will determine if there is an association between testing positive for hepatitis C virus and breast cancer while controlling for age. Both Hepatitis C and breast cancer are extremely prevalent in Egypt (?). Very 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 a area of interest (?). Thus, research surrounding these diseases is becoming more important in the field.

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

## 2 Data

The data used in the paper is from a retrospective case control study conducted in 2020 (?). The data includes 405 subjects as part of the study group, all having been treated for breast cancer in the past 10 years. The second group consists of 145 adult females from a governorate in Egypt, who all participated in a cross-sectional study from 2015-2017. This data can be put into a 2x2 contingency table for statistical analysis.

## 3 Methods

To test the association of breast cancer and Hepatitis C, one can use the chi-square or Fisher's Exact test, depending on the conditions from the sample size (?). 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. As a result, we can determine if women who chronically test positive should be tested more frequently because they are at a higher risk.

Figure ?? shows the probability of the original table occurring.

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

Figure 1: Hypergeometric pmf.

## 4 Application

2x2 Contingency Table for all patients:

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

2x2 Contingency Table for patients younger than 45:

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

2x2 Contingency Table for patients older than 45:

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

## 5 Discussion

-