



Predicting overall survivability in comorbidity of cancers: A data mining approach



Hamed Majidi Zolbanin*, Dursun Delen, Amir Hassan Zadeh

Department of Management Science and Information Systems Spears School of Business, Oklahoma State University, Stillwater, OK, United States

ARTICLE INFO

Article history:

Received 7 February 2015

Received in revised form 1 April 2015

Accepted 2 April 2015

Available online 19 April 2015

Keywords:

Medical decision making

Comorbidity

Concurrent diseases

Concomitant diseases

Predictive modeling

Random forest

ABSTRACT

Cancer and other chronic diseases have constituted (and will do so at an increasing pace) a significant portion of healthcare costs in the United States in recent years. Although prior research has shown that diagnostic and treatment recommendations might be altered based on the severity of comorbidities, chronic diseases are still being investigated in isolation from one another in most cases. To illustrate the significance of concurrent chronic diseases in the course of treatment, this study uses SEER's cancer data to create two comorbid data sets: one for breast and female genital cancers and another for prostate and urinal cancers. Several popular machine learning techniques are then applied to the resultant data sets to build predictive models. Comparison of the results shows that having more information about comorbid conditions of patients can improve models' predictive power, which in turn, can help practitioners make better diagnostic and treatment decisions. Therefore, proper identification, recording, and use of patients' comorbidity status can potentially lower treatment costs and ease the healthcare related economic challenges.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Cancer is the second leading cause of death in the United States.¹ It is also a major cause of death worldwide, especially (and ironically) in high income countries.² Research on causes and behavior of cancers has resulted in significant advances in our understanding of the disease over the past four decades. Even though cancer studies have traditionally been clinical and biological in nature, the recent technological advances have made data driven analytic studies a common complement.

The exploration of massive medical databases with the aid of new computational tools has confirmed the existence of coexisting diseases, including certain cancers. However, current medical research has a tendency to follow a reductionist approach to the study of ailments by investigating them in isolation from one another, rather than considering their interactions [57]. Recent findings urge taking a different stance toward comorbid diseases by denoting how coexisting illnesses might affect the diagnosis, treatment, and evaluation of treatment effectiveness, as well as survival of patients [2,20,21,26,27,45,61].

Yet, another equally important reason for the consideration of comorbidities is their impact on treatment costs, which in turn affect

economies. According to the National Health Council, 133 million Americans are affected by incurable, ongoing chronic diseases, and this number is expected to grow to 157 million in 2020, with 81 million suffering from multiple conditions.³ These figures find more salience with chronic conditions accounting for more than 75% of all healthcare costs. While in 2007, \$1.3 trillion was reported as the adverse economic impact of chronic diseases, including cancer, it is projected to increase to \$4.2 trillion for superfluous treatment costs and lost economic output.

Prior research has shown that cancer treatment recommendations might significantly be altered based on the severity of comorbidities. Specifically, the extent of the tumor spread is not the sole indicator of treatment. Instead, the overall health of the patient might have a greater weight in choosing the treatments [17]. Regarding the serious interplay of coexisting complications with all different phases of cancer treatment, and with the increasing trend in the development of intercurrent illnesses, the present cancer classification system needs to be revised, as it does not account for the severity of comorbid conditions [46]. Even if concurrent health issues are diagnosed and accounted for during the course of treatments, excluding them from general data sets or storing them in disconnected systems hampers prospective statistical analyses that might reveal useful patterns about their interplay. Similarly, elimination of comorbidity information may compromise the effectiveness of clinical decision support systems (CDSS). These systems “apply best-

* Corresponding author.

E-mail address: zolban@okstate.edu (H.M. Zolbanin).

¹ <http://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm>.

² <http://www.who.int/mediacentre/factsheets/fs310/en/index1.html>.

³ http://www.nationalhealthcouncil.org/NHC_Files/Pdf_Files/AboutChronicDisease.pdf.

Download English Version:

<https://daneshyari.com/en/article/554702>

Download Persian Version:

<https://daneshyari.com/article/554702>

[Daneshyari.com](https://daneshyari.com)