

Accepted Manuscript

Temporal Electronic phenotyping by mining Careflows of Breast Cancer Patients

A. Dagliati, L. Sacchi, A. Zambelli, V. Tibollo, L. Pavesi, J.H. Holmes, R. Bellazzi

PII: S1532-0464(16)30186-1

DOI: <http://dx.doi.org/10.1016/j.jbi.2016.12.012>

Reference: YJBIN 2699

To appear in: *Journal of Biomedical Informatics*

Received Date: 8 July 2016

Revised Date: 21 November 2016

Accepted Date: 31 December 2016

Please cite this article as: Dagliati, A., Sacchi, L., Zambelli, A., Tibollo, V., Pavesi, L., Holmes, J.H., Bellazzi, R., Temporal Electronic phenotyping by mining Careflows of Breast Cancer Patients, *Journal of Biomedical Informatics* (2017), doi: <http://dx.doi.org/10.1016/j.jbi.2016.12.012>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Temporal Electronic phenotyping by mining Careflows of Breast Cancer Patients.

A. Dagliati¹, L. Sacchi¹, A. Zambelli^{2,3}, V. Tibollo², L. Pavesi², J.H. Holmes⁴, R. Bellazzi^{1,2}

¹Department of Electrical, Computer and Biomedical Engineering and Centre for Health Technologies, University of Pavia, Italy

²IRCCS Fondazione S. Maugeri, Pavia, Italy

³Ospedale Papa Giovanni XXIII, Bergamo, Italy

⁴ Department of Biostatistics and Epidemiology, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

ABSTRACT

In this work we present a careflow mining approach designed to analyze heterogeneous longitudinal data and to identify phenotypes in a patient cohort. The main idea underlying our approach is to combine methods derived from sequential pattern mining and temporal data mining to derive frequent healthcare histories (careflows) in a population of patients. This approach was applied to an integrated data repository containing clinical and administrative data of more than 4,000 breast cancer patients. We used the mined histories to identify sub-cohorts of patients grouped according to healthcare activities pathways, then we characterized these sub-cohorts with clinical data. In this way, we were able to perform temporal electronic phenotyping of electronic health records (EHR) data.

Key Words: Electronic phenotyping, Careflow mining, Temporal data mining, Heterogeneous data sets

A. INTRODUCTION

During disease progression, patients undergo numerous events such as hospital admissions and discharges, lab tests, and follow-up visits. These events occur in a temporal sequence that represents the *patient careflow* [1], [2]. Their automated identification from clinical data still represents a major unmet clinical need. As a matter of fact, clinical guidelines and care protocols have been identified as means to improve and standardize health care services, but in the absence of an effective Information and Communications Technology (ICT)-based solution to automatically extract frequent careflows, it is often impossible to measure their implementation. Moreover, the identification of the careflows that occur in clinical care allows profiling both the activity of health care providers and to have insights on the groups of patients actually treated.

Download English Version:

<https://daneshyari.com/en/article/4967026>

Download Persian Version:

<https://daneshyari.com/article/4967026>

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