Process Mining and Clinical Pathways: an application to Breast cancer data in Tuscany

Francesca Ferrè[1], Chiara Seghieri[1], Andrea Burattin[2], and Andrea Vandin[1,2]

[1] Sant'Anna School of Advanced Studies, Pisa, Italy
[2] DTU Technical University of Denmark, Kgs. Lyngby, Denmark

In the last decades there has been a significant increase in the use of clinical pathways in developed countries. A clinical pathway is an instrument defining the optimal sequence of clinical interventions, timeframes and expected outcomes for selected homogeneous groups of patients. It can be viewed as an evidence-based standardized treatment process aiming at disseminating among the healthcare professionals the adoption of best-practice clinical standards to improve quality and equity of care while reducing fragmentation, and ultimately costs.

An example are pathways used to improve care for cancer patients [1]. Cancer units and care networks are developing or using pathways for improving clinical governance, reducing unwarranted variation in quality of care and inequities in service provision (i.e., differences that can be considered as "avoidable" or "unjust"). Specific cancer care pathways are set for the different cancer pathology and often are designed by clinical teams to suit local requirements. Despite the growing popularity of clinical pathways there is still lack of studies recording and analyzing potential deviations of practice from them. The clinical pathway should be designed for the needs of a "standard" patient, therefore "positive" variations might occur for patients with specific needs. On the other hand, we should be able to identify negative variations that are not justified and might be attributable to system inefficiencies, clinical behaviour, supply side-induced demand, etc. To date, most studies are focused on assessing the effectiveness of such pathways, specifically clinical changes achieved and change in patient satisfaction [3].

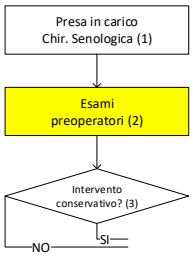

Fig. 1. Exc. from [1].

We aim at studying how process mining methods, already successfully applied to the health domain (see, e.g., [2]), can help in studying breast cancer pathways, e.g. to: *(i)* identify common and uncommon paths of patients with breast cancer, or to *(ii)* compare what happens in practice (the *mined pathways*) with pathway recommendations (the *theoretical pathways*) to identify potential unwarranted variations and understand the drivers of these variations.

We study the breast disease care pathway depicted in Fig. 1 introduced in 2021 by Regione Toscana for the cancer care network running in Tuscany, Italy (the picture displays just the initial steps). The pathway focuses on the flow-chart of the surgical treatments following the discovery of a breast lump (box 1, not including adjuvant therapies for the treatment of breast cancer), followed by pre-operative examinations (box 2) which might lead to conservative

or non-conservative surgeries (respectively denoted by the choices *SI* and *NO* of diamond 3). The pathway includes further steps not discussed here for brevity.

We used administrative data on all examinations and surgeries in Tuscany for the years 2018-2019. We selected breast cancer patients undergoing surgical treatment (conservative and non-conservative surgery) in all public hospitals in 2019 and linked the data about pre-operative examinations. We further pre-processed the data in Python to isolate female single-surgery patients. Examination data contained two dates: the one of reservation of the exam, and that of the exam, considered as two separate events. Likewise, surgeries contained four dates: the one in which the patient entered the waiting list, and those of beginning and end of the hospitalization, and of the surgery. We transformed each such date in a separate event. Overall, the cleaned data regards about 2 300 patients for about 50 000 events. This richness of information allows for several analyses. Fig. 2 provides one regarding waiting times. Timely access to care is key for an effective and satisfactory care pathway. For instance, 30 days should be the maximum waiting times from breast cancer diagnosis to surgery. We studied the flows for conservative (Fig. 2 left) and non-conservative (right) surgeries. Starting from the top, the two flows contain boxes for: entry in the waiting list, begin of hospitalization, surgery, end of hospitalization. Indeed, Fig. 2 has been obtained using Fluxicon Disco by focusing on these four events and distinguishing among the two surgery types. The arrows labels denote the median time elapsed and the absolute frequency. From the top red arrows we can see that the 30 days requirement is satisfied in most of the cases for both surgery types. Furthermore, as expected, median hospitalization length is one day for conservative surgery, while for non-conservative ones it is about two days as they include reconstructive surgery.

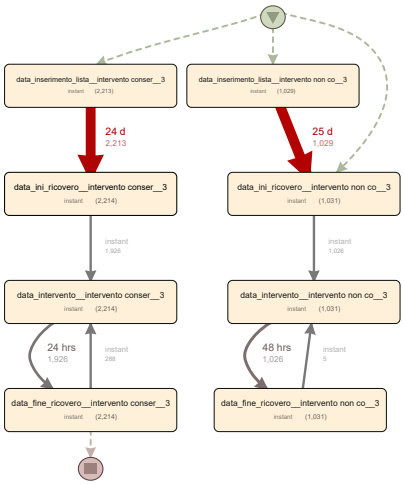

Fig. 2. Preliminary results.

These preliminary results show the potential of process mining techniques in studying the level of adherence to evidence-based clinical guidelines, and motivates in pursuing this intriguing research line, especially in the post-COVID era where health processes need to become more efficient and more inclusive.

Acknowledgments Partially supported by DFF project REDUCTO 9040-00224B

References

1. Decree 2987/2021 of Regione Toscana, Annex A. `https://bit.ly/pathwayBC`
2. Fernandez-Llatas, C.: Interactive Process Mining in Healthcare. Springer (2020)
3. Nuti, S., Ferré, F., Seghieri, C., Foresi, E., Stukel, T.A.: Managing the performance of general practitioners and specialists referral networks: A system for evaluating the heart failure pathway. Health Policy **124**(1), 44–51 (2020)