Exploiting Natural Language Processing for Improving Health Processes

M. van Keulen[1], J. Geerdink[2], G. Linssen[2], R.H.J. Slart[1], and O. Vijlbrief[2]

[1] University of Twente, {m.vankeulen;r.h.j.a.slart}@utwente.nl
[2] Hospital Group Twente, {j.geerdink;g.linssen;o.vijlbrief}@zgt.nl

1 Introduction

In the medical world, high quality digital registration in an Electronic Patient Dossier (EPD) of symptoms, diagnoses, treatments, test results, images, interpretations, and outcomes becomes commonplace. Together with a shortage of medical professionals, means that they experience pressure at the expense of actual 'hands on the bed'. On the other hand, EPDs contain a wealth of largely unused, unstructured textual information. Clinicians primarily communicate with each other through letters and reports. Our main question is: Can Natural Language Processing (NLP) exploit this wealth? By extracting structured data and using it as features for machine learning, a wide variety of process improvements become possible. Furthermore, it may contribute to the desire of government and health stakeholders to simplify registration and relieve pressure. This paper sketches a few prominent process improvements that we plan to research.

2 Potential ways of improving health care processes

A wealth of textual data in EPDs. A care process often consists of many steps involving many actors (see Figure 1(a)). A visit to a general practitioner (GP) often results in a referral to a specialist by means of a letter containing observations and suspicions. At every stage, clinicians document observations and interpretations of consults, tests, images, treatments, and operations in reports.

Improving Complex Diagnostic Processes. Many conditions are hard to diagnose. For example, Giant Cell Arteritis (GCA) is an immune mediated vasculitis characterized by inflammation of the large and medium sized arteries. It is uncommon (25–70 per 100.000) and presents itself only with atypical signs and symptoms such as low-grade fever, malaise, and weight loss. Its diagnosis is difficult and can be missed easily. Untreated it may lead to loss of eye sight and cerebral stroke. Several diagnostic modalities are essential for diagnosis, but machine learning may pick up subtle patterns of clues scattered over various reports containing conclusions from these modalities. Faster referral from primary to secondary care is beneficial [1]. By varying the *data window* (see Figure 1(a)) and assessing prediction performance, one may determine from which step in the care process a sufficient prediction performance may be expected. A learned

score	malignant
1 & 2	0%
3	0-2%
4a	2-10%
4b	10-50%
4c	50-95%
5	>95%
6	proven
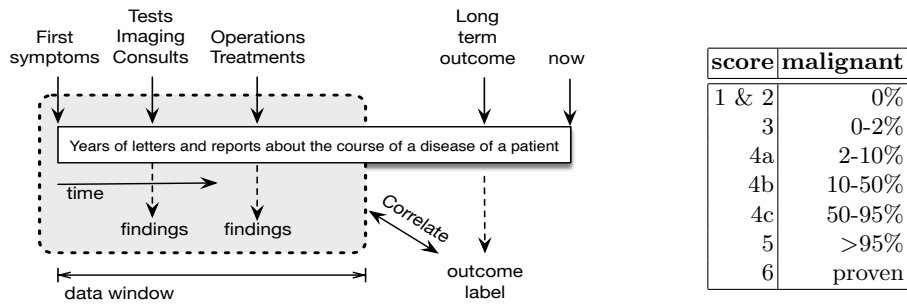
(a) Different data windows have different prediction performance   (b) BIRADS score

model may be unobtrusively incorporated in existing EPD software warning the clinician of substantial risk for uncommon conditions like GCA.

As another example, chemo and radiation therapy in cancer patients damage heart and bloodvessels. It is important to identify those patients who are at high risk for long term complications. They need tailored medical therapy to prevent cardiovascular diseases. Machine learning may aid in identification, avoiding or suggesting such measures [2].

Dynamic Shortening of Care Processes. Actions like tests, imaging, treatments, etc. are preceded by consults. Patterns may be discovered between data in letters and reports from before a consult, that sufficiently predict a subsequent action. In such cases, one may dynamically adapt the process to already plan this action before the consult, at the limited expense of some cancellations due to mispredictions, or one may even decide to omit the consult altogether.

Quality assurance. Hospitals often enforce strict protocols as a means for quality assurance. NLP techniques may aid in checking adherence to protocols and supporting clinicians in upholding the standards. For example, the radiology diagnostic reports for breast cancer in the Hospital Group Twente contain among other things a standard risk classification, called "Breast Imaging Reporting and Data System" (BIRADS; see Figure 1(b)). Whether or not these risk scores correlate well with actual outcomes is largely unknown. Being able to check this is important for quality assurance. Moreover, discovering under which circumstances deviating risk assessments are given, may aid in improving care processes with better-informed decisions [3].

References

1. Prior, J., Ranjbar, H., Belcher, J., Mackie, S., Helliwell, T., Liddle, J., CD, C.M.: Diagnostic delay for giant cell arteritis - a systematic review and meta-analysis. BMC Med. **15**(1) (June 2017) 120
2. Rumsfeld, J., Joynt, K., Maddox, T.: Big data analytics to improve cardiovascular care: promise and challenges. Nat Rev Cardiol. **13** (2016) 350–359
3. Sippo, D., Warden, G.: Automated extraction of bi-rads final assessment categories from radiology reports with natural language processing. J Digit Imaging **26**(5) (October 2013) 989–94