Biomarkers in non-small cell lung cancer

Societal Impact and Valorisation

Sara Carvalho

SOCIETAL IMPACT AND VALORISATION

In this thesis, processes of data gathering, predictive and prognostic inferences derivation from non-invasive sources of information are thoroughly described. These approaches were discussed in the context of a specific oncological disease – lung cancer, but should not be seen as restricted to that, instead as a powerful way of tackling others, as for instance head and neck, rectal or cervical cancer. In this section the possibility of creating value from the knowledge here presented will be discussed, as well as the approaches that could be followed in order to positively impact people's treatments and potentially improve outcomes.

We face these days a great increase in data availability. This "big data" era is opening branches for data collection, analysis and exploration that will definitely shape the way life is known. Medicine is no exception to this trend. On one hand, the large amount of data and its immediate availability requires new techniques for analysis. This inherently leads to a better insight of its underlying information, going beyond classical statistical techniques. More powerful predictive tools are becoming available, enhancing outcome prediction based on the existent sources, by means of using, for instance, machine learning techniques. This is aligned with the aimed personalized care – bringing the right treatment to the right patient. The use of prognostic models is increasing, with many online tools already available:

- http://www.predict.nhs.uk/ is an online tool used to help deciding the ideal course of treatment following breast cancer surgery, based on one's cancer histopathology;
- http://www.cancer.gov/bcrisktool/ allows estimating woman's risk of developing invasive breast cancer;
- https://www.mskcc.org/nomograms a long list of nomograms developed for a great variety of solid tumours, and based on hundreds or even thousands of patients, can be used to predict cancer outcomes or assess disease risk;
- http://www.predictcancer.org/ a platform making available published models developed at Maastro for lung, rectum, head and neck, and endometrium cancer

In this last one, in due time all the models presented in this thesis will be made available for risk probability assessment of individual patients. This shows one of the greatest outcomes of this research – knowledge sharing. This is further complemented with the sharing of data used for models development and validation at https://www.cancerdata.org/. We believe that great gain is obtained in data sharing. First, this will enable others to assess reproducibility, and allows the investigation of other methodologies and hypothesis in same data. Secondly, making data available will improve other institutes validation procedures, and will help overcoming the main drawback of such data-driven approaches: the lack of enough data for both development and validation purposes. Finally, collec-

tion and standardization of data will be stimulated to validate previously developed models which will also make it possible to update and improve those.

After this general remark, valorisation of the knowledge derived from these two non-invasive sources of information will be further discussed referring to each independently.

PART 1. IMAGING BIOMARKERS

Comprising the largest segment of the research conducted in the scope of this thesis, imaging biomarkers were devoted 4 chapters in the present work. Main hypothesis behind the complex imaging analysis here explored followed a single premise: more information is nowadays available in medical images, than circumscribed to most common imaging features, popular in published literature, such as tumour volume or maximum intensity measured with FDG uptake in PET scans. Radiomics offers a high throughput of imaging descriptors, capable of characterizing particular regions of interest (commonly tumours, in the Oncology field), by means of shape and size features, first order statistics - regarding pixel values distribution on corresponding histograms, and texture information, based on the arrangement of consecutive and neighbouring pixels.

Standard practice imaging is commonly used for treatment design and dose planning in radiotherapy facilities. In this project it was investigated how imaging can potentially be even more influential on treatments. The use of imaging for precision medicine instead of a one size fits all approach may be a reality within reach following the positive preliminary conclusions presented along this thesis. Nowadays picture archiving and communication systems (PACS) connect images to healthcare systems (Electronic Medical Records - EMR), that ultimately open the branch for integration with clinical decision apps, providing a fast service anywhere, anytime.

With Radiomics one can benefit from the re-use of an existing source of information: the non-invasive character and promise of a fast and cost effective diagnostic tool will always grant imaging to be a reference tool for disease management. In the end patients will in almost all cases be scanned for diagnosis, treatment planning and follow-up purposes. Radiomics is aligned with these data exploiting we are witnessing nowadays, by converting data into information, based on an existing source: imaging.

During the past years, research has been conducted at Maastro and other institutes in different continents, in an attempt to highlight the advantageous use of medical imaging and its thorough investigation. Ever since, the Radiomics approach here explored has resulted in multiple publications (scientific papers, dissertations, etc.), that show that the presented work has great application possibilities. Currently making its transition into a software application for imaging analysis, Radiomics toolbox will soon become available for purchase, sponsored by OncoRadiomics, a spin-off company that made its way from an idea up to a final product - RADIOMICS™ (http://www.oncoradiomics.com/).

Strategies for certification are necessary (within Europe (CE marked) and USA (FDA approved)), so that such software could be made available as a plug-in to be integrated in several software packages. Development was performed in a modular way to allow an easy incorporation into treatment planning and diagnostic radiology software from medical companies (RaySearch, Varian, Aquilab, Siemens, Philips, Intuitim and Elekta, for instance). Additionally, it could be made available as software as a service (SaaS) as to provide means for patient stratification and response assessment in pharma-sponsored clinical trials. Following preliminary evidences demonstrated in chapter 6 of this thesis, applicability of RADIOMICS™ to monitor longitudinal treatment response, is further warranted for DeltaRadiomics™ applications.

Application of RADIOMICS™ to other cancer sites will be investigated in collaboration with renowned clinical centres. Marketing strategy for this encompasses publication of peer-reviewed articles and presentations at key conferences. Currently, several world-leading cancer centres are using a research version of RADIOMICS™ to support its advantageous use, by generating evidence and awareness of mentioned technology.

PART 2. BLOOD BIOMARKERS

With biopsies not being a possibility for many lung neoplasms, the possibility of inferring tumour's microenvironment from circulating oncoproteins is a powerful alternative. In this thesis we based our blood biomarkers selection to those related to known adverse and intrinsic characteristics of lung tumours: hypoxia, inflammation, tumour load and immunogenic descriptors. This pre selection of markers was based on literature evidence of their prognostic value, documented on at least 100 patients. We made available for the medical community, anonymized patient data comprising almost 400 of patients with up to 12 dissimilar markers (https://www.cancerdata.org/resource/doi:10.17195/candat.2016.04.1). We expect with this to engage more institutes in performing similar approaches, testing their hypothesis and improving outcomes using blood biomarkers.

Currently, new methodologies are being developed as, e.g. multiplex panel, allowing multiple biomarkers to be measured simultaneously, even ones present in low abundance. These technologies still need to be validated and standardized before introduction in clinical practice but their quick development opens possibilities for, not only, baseline assessment, but also monitoring during and after treatment in a fast and convenient way.

As a general conclusion, two fundamental steps are to be followed in order to bring the research performed to a higher level of beneficial gain to the society:

- Raising awareness on standardized data collection and sharing, will greatly improve the way Medicine is done nowadays, with unlimited benefit expected to be re-

trieved from several sources of information. This increase on data availability, will as well help develop a superb variety of approaches on data analysis, following different hypothesis.

- Engage more institutes in using Radiomics software implementations, in its research version as to generate evidence of its great usage; and later by its commercialization and integration with medical software platforms.