Image features for the future in stage IV non-small cell lung cancer

Valorisation addendum

Prognostic models

Although there have been innovations in treatment options for stage IV non-small cell lung cancer, the overall survival rate for these patients is still low. Innovations in treatment consist for example of specific drugs that are optimized for various cancer types, e.g. EGFR mutated tumors treated using tyrosine kinase inhibitors. However, this treatment is not effective in, for example, patients with a KRAS mutation. And if patients do not respond on a certain treatment, quality of life of these patients can be improved by stopping the treatment. In the first part of this thesis, prognostic models including image biomarkers are developed and validated.

Prognostic models help in selecting patients that benefit from certain therapies by which side effects are reduced and quality of life is improved. If a tumor is not responding on a specific chemotherapy, another treatment may be used which is potentially more effective. Prognostic models already are able to predict the chance that a patient will respond on the treatment before the start of treatment. We showed that PET-based response assessment showed response already after three weeks of treatment, which is earlier than CT-based response assessment usually performed after 6 weeks of treatment. Early response assessment and prognostic models will lead to economically benefits through less side effects, less expensive treatments and better prognosis. Additionally cost effectiveness analyses should be performed in which the relative value of a specific treatment is measured as the additional cost to achieve an incremental health benefit. Information from prognostic and/or predictive models can be used as input for cost effectiveness analyses to predict survival or side effects, by which life expectancy adjusted for morbidity or quality of life for different treatments can be compared.

Next to that, prognostic models can be integrated in decision support systems. Decision support systems help patients together with their physician to choose between different treatments, based on results of previous studies and personal preferences. For some patients, quality of life is more important while others prefer a prolonged survival. On www.predictcancer.org prognostic models for overall survival, local recurrence, cost effectiveness and side effects like dysphagia can be found for brain metastases, endometrium cancer, head and neck cancer, lung cancer, esophageal cancer, prostate cancer and rectum cancer. With initiatives like predictcancer.org decision support systems may easily be implemented in the clinic.

Image features

Imaging is an important part of routine care in diagnosis, treatment and follow up in oncology. Image biomarkers improve prognostic models in a non-invasive way. Image biomarkers can be qualitative, so called semantic features, but also quantitative, so called radiomic features. While scoring semantic features is time consuming and inter observer dependent, radiomic features are automatically extracted on repeated imaging. Imaging is non-invasive and widely available because it is used in standard clinical care. In chapter 8 we showed that there are associations between semantic features and radiomic features that are correlated to mutation status of the tumor. When imaging is used to determine the mutation status of tumors, invasive biopsies that are only samples of mostly heterogeneous tumors can be reduced or in the future are not needed anymore.

Many publications already showed the potential of radiomics. Radiomic signatures can often be used in different tumor sites. Software developments are ongoing and more and more open source radiomic packages become available. By making radiomic software open source, usage by more research departments makes external validation of radiomic signatures much easier. Next to that, initiatives are ongoing in which radiomics is implemented in clinical software. An example of this is the collaboration between OncoRadiomics and AQUILAB in the ARTIVIEW software package. Such software could be used in either clinical studies, to develop new predictive signatures or in clinical decision support systems based on defined signatures. These commercial

software packages consist of a database system in which multimodality images could be saved, fused or registered, segmented and used for dose calculations. By implementing radiomics in already clinical used software packages, radiomics can easily be used in diagnosis, segmentation and response assessment. External validation studies validated existing radiomic signatures in different tumor sites as well as using different image settings. Radiomic signatures that are externally validated in large datasets will lead to useful clinical products.

Radiomics may also be very useful in diagnosis. In chapter 9 we showed that differences in delta texture radiomic features over time might reflect myosteatosis. This shows that using radiomics biological processes behind some diseases, for example cachexia can be unraveled. This can in the future possibly lead to new treatment options. An increasing number of studies is looking into radiomics for differentiating benign lesions and malignant lesions in for example breast, prostate and lung. Using this information leads to earlier diagnosis and therefore a prolonged survival and better quality of life.

The most time consuming step in the radiomics process is still segmentation of the region of interest. Machine learning, which is a subset of artificial intelligence to automatically detect patterns in data, by which segmentation is automated, combined with radiomics, will in the future automatically analyze large datasets used for prognostic or diagnostic models.

Image quality

Although for automatic extracted features it is important that the images used are acquired on a standardized way. In the second part of the thesis image quality in multicenter trials is reviewed and from that recommendations are formed to improve image quality in future multicenter trails. When using images from multiple centers there is a large variation in image acquisition and reconstruction settings, which influence radiomic feature values and feature stability and therefore results from prognostic models including radiomic features. In chapter 5 we showed that slice thickness influences radiomic features, however this can be reduced by resampling the images prior to feature extraction. Studies investigating the influence of scan parameters on radiomic features will lead to a more standardized method for imaging in radiomic studies and therefore in less variation in radiomic features and more reliable radiomic signatures.

Although guidelines for PET imaging in multicenter trials exist, in chapter 3 we showed that there is still a large variability in image acquisition and reconstruction parameters between Dutch institutes, which is probably even larger between European centers and centers worldwide. The largest difference was seen in the time between FDG injection and the start of the PET scan. Longer uptake times will lead to higher SUV values, which for response assessment can lead to a false response assessment. When using PET scans from multiple centers in response assessment studies we recommend to use a prospective central quality review to reduce the variability between institutes and to better comply with the existing guidelines, which leads to a better overall image quality, higher reproducibility and less inter-center variation.

For contrast-enhanced (CE-)CT in diagnosing asymptomatic brain metastases in non-small cell lung cancer no clear guidelines exist. CE-CT scans with the arms of the patients next to the head are often reported to be diagnostic, while the arms of the patients induce beam hardening which influences the image quality. In chapter 4 we showed that it is important that clinicians inform imaging experts about the indication of a specific scan to help radiologists choose the optimal scan protocol and to check the image quality for the requested purpose. This will lead to a more uniform interpretation, less variation in multicenter trials and uniform treatment decisions; all improving treatment quality.