Precision medicine and quantitative imaging in glioblastoma

ELMED219 (Artificial Intelligence and Computational Medicine), January 3-28, 2022

https://github.com/MMIV-ML/ELMED219-2022

Team #5

1 Research plan

1.1 A brief background to the field

Glioblastoma (GBM; grade IV astrocytoma), is the most common and most aggressive malignant primary brain tumor in adults. It accounts for 48.3% of all malignant brain tumors [1], and has one of the highest mortality rates of all malignant tumors with a median survival rate of 12 to 15 months after diagnosis. Compared to other low grade gliomas the growth is invasive. This makes treatment challenging and early diagnosis is crucial for a better prognosis.

Standard procedure is to use molecular markers to decide treatment and to evaluate glioblastoma. eIsocitrate Dehydrogenase 1 and 2 (IDH1/2) are two of these markers and have become very important for how GBM. In 2016 WHO incorporated the mutational status of these molecular markers into the classification of glioblastoma. Mutations in these where found in over 70% of lower grade gliomas and in secondary glioblastomas, which evolves from lower grade glioma. [2].

Methyliation of O(6)Methylguanine-DNA Methyltransferase (MGMT) is another important biomarker, and it gives an indication on how receptive a patient will be to chemotherapy. Temozolamide (TMZ) is the main chemotherapy for GBM and is a DNA alkylating agent. Resistance to TMZ can be mediated by the MGMT DNA repair enzyme and with methylation of MGMT, it will silence this gene. For this reason, MGMT methylated patients are more receptive to chemotherapy[3]. In addition to the previously mentioned genetic biomarkers, 1q19p-codeletion and H3K are also important for diagnosing GBM according to WHO [4].

Misdiagnosis of the tumor affects the medical intervention and reduces the chances of survival of patients [5]. MRI provides a massive amount of information to the clinicians. However, the clinicians are typically restricted to qualitative descriptors or subjective quantitative assessments to articulate changes in imaging. The resulting clinical evaluations have significant potential for bias [6]. To further get information about the genotypes of the tumor, to differentiate the diagnosis in high-grade or low-grade glioblastoma and to predict the efficiency of systemic agents, it requires tissue from the tumor which introduces the patient to invasive methods [6]. The procedure of taking brain biopsies is time and resource demanding, expensive and is associated with risk and discomfort or pain for the patient. Further, patients diagnosed with GBM

often get reduced life quality due to symptoms and side effects of the therapy, and as the question how long the patient is expected to live, is raised. Survival prediction is of vital importance to clinicians, patients, and their families [7], and most clinicians overestimates cancer patients' survival [8]. To our knowledge, several parameters within the patient and tumor features has shown to affect the overall survival for GBM patients [9, 10, 11, 12, 13, 4] .

1.2 Objectives and expected impact

Both the patients and the caregivers describe the need for hope, support and information following the diagnosis of GBM [7]. In the aim of providing precision medicine, genetic information from the patient is inevitable. The four genetic biomarkers that WHO has established after several studies shows associations between genotype and overall survival [4]. However, today's methods to provide genotypes is invasive and overall survival rate is often hard to predict.

Based on non-invasive methods and machine learning we aim to predict the overall survival (OS; in months) based on the features of the tumor, genotype, demographic data and clinical data. We hope the method of deciding genotype and prediction of survival becomes an easily accessible, non-invasive, and a supportive tool for the caregiver to decide the best treatment option. For the patients and their families, we hope that the answer of how long the patients have left will increase the quality of life in the, probably, last months of patients' lives. To do this, we aim to use radiomics in synergy with liquid biopsies to provide genetic information together with clinical parameters to predict the patients overall survival.

1.3 Material and methods

Figure 1 is a schematic review of the process of creating our model that can predict OS for patients with GBM. The first input is original data which is used to produce synthetic data by Generative Adversarial Networks (GANs). The synthetic MRI images are put into a deep learning model for feature extraction that leads to genetic biomarkers and then these are used for classification of the tumor. The entire data set including clinical data and genetic biomarkers are put into a regression model to predict OS.

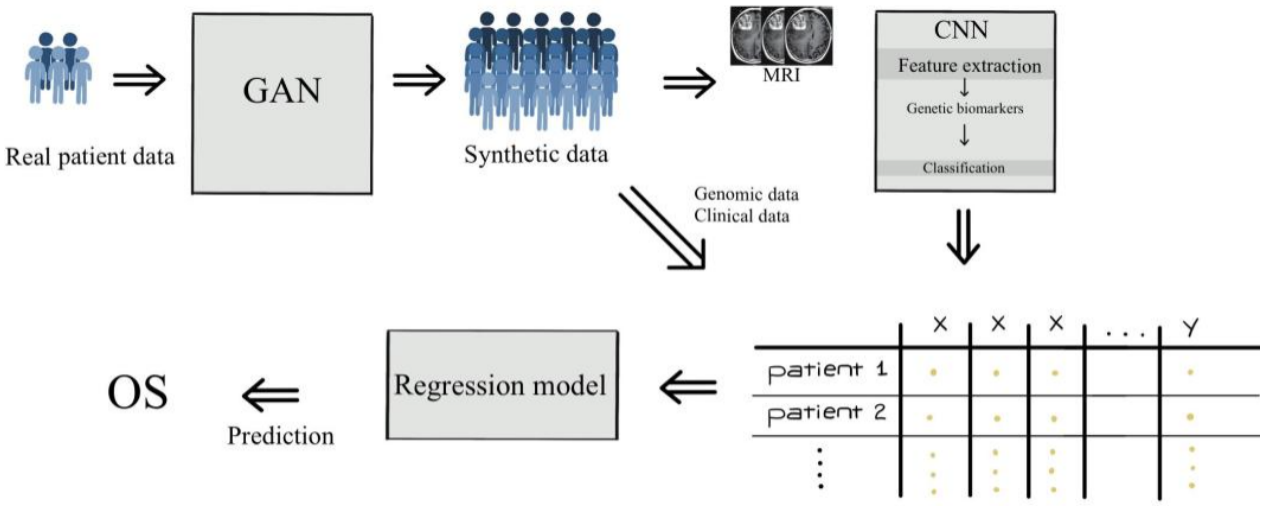

Figure 1: A schematic review of the process of creating our model that predicts OS for patients with GBM (made by Caroline Haugen).

1.3.1 Radiomics

Magnetic resonance imaging (MRI) is the standard procedure for all GBM-patients and therefore MRI data will always be available for an oncologist. We want to use this already available data as input for our model. MRI is a non-invasive imaging technique, where the patients body is surrounded by a strong magnetic field that aligns the proton spins in the patients body. The MRI pictures accumulate an amount of information that expands beyond the limitations of human eyes capability to study and interpret the pictures. A machine is able to extract a large number of quantitative features from the MRI picture. The features are for example textural features, which are obtained by voxel to neighbor voxel relationships. An example of textural features are those derived from the grey level co-occurence matrix (GLCM), which quantifies within a given range the number of voxels with the same intensities. The machine can be trained to classify the genotype of the tumor based on these quantitative features. This practice of extracting genomic information from radiologic data, is called radiomics. We will use a deep learning model to extract the features. Our model will be able to learn thousands of features from an image without the need of segmentation of the volume of interest (VOI) by a radiologist. After extracting the features, the model will perform multi-label classification of the tumor according to the four genomic biomarkers. And finally our model will use these biomarkers together with the other features to predict OS.

Deep learning is a form of Artificial Intelligence that works by trying to imitate how humans think by analyzing patterns in large data sets. Our model uses a convolutional neural network (CNN) which works like neurons in the human brain. The neurons work together to detect characteristics of an input and develop an output with the processed image [14]

The CNN consists of multiple layers (or clusters), with each level focusing on tasks between a low level feature or a high level feature. The different clusters are able to communicate and share information with each other as they work. The first convolutional layer is called the kernel. The kernel shifts over the image and performs matrix multiplication between the kernel and the portion of the image the kernel hovers over. The kernel moves over the entire picture until it reaches the end of the image. The first layer is to extract low-level features, such as edges, but with added layers, we are also getting the high-level features that gives us a whole understanding of the image.

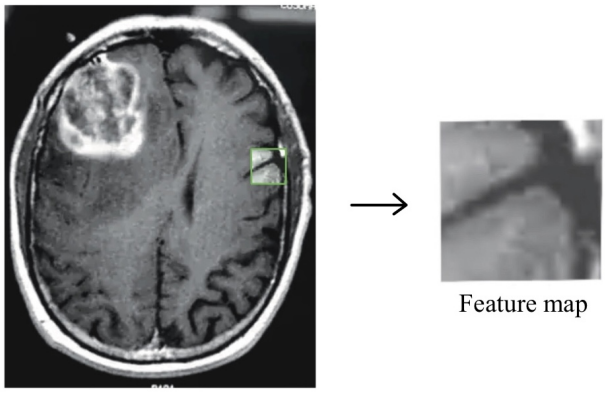

Figure 2: The convolutional layer finding low level features in a MRI scan of a brain with GBM

There is also a pooling layer that has the task of reducing the spatial size of the convolved feature. This is to minimize the computational power used to analyze the data. It also extracts dominant features. A large number of features per a low number of samples imposes the risk of overfitting. Performing a dimension reduction of our extracted features are crucial to obtain a generalizable model. There are two types of

pooling: max pooling and average pooling. Max pooling returns the maximum value of the portion of the image the kernel is covering, whereas average pooling returns the average value. For our project we will use max pooling. The convolutional layer and the pooling layer together form a layer of the convolutional network. After the process is gone through, we now have a model that is able to understand the features of an input image [15].

To increase reproducibilty between different datasets some image processing is done before the model extracts the features. We will use the library pyRadiomics for this purpose. This preprocessing includes interpolation to isotropic voxel spacing, normalization and discretization. Normalization is performed to remove voxels from the segmentation that fall outside a range of grey levels defined as plus minus three standard deviations from the mean gray level. Discretization of image intensities inside the VOI consists in grouping the original values into bins and it is essential to make feature calculation tractable.[16] For the feature extraction we will adhere the guidelines of the Image Biomarker Standardization Initative (IBSI). These guidelines offer a consensus for standardized feature calculations. In Figure 2 the workflow of radiomics is summarized.

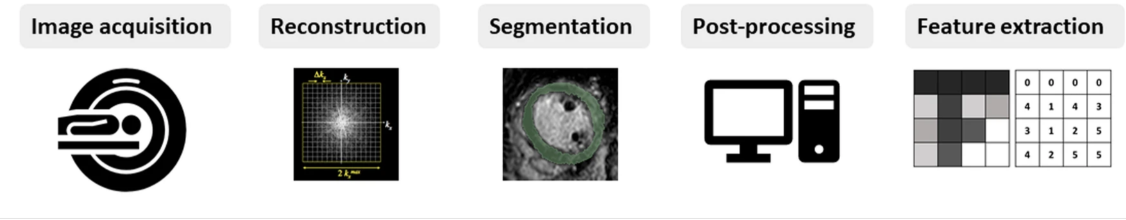

Figure 3: The workflow of radiomics [16]

.

1.3.2 Liquid biopsies (LB)

Similarly to MRI, blood samples are also taken as a standard procedure in diagnosing GBM-patients. Current diagnosis of this disease typically relies on imaging techniques and tissue biopsies, which have their inherent limitation being difficulties in detecting the true tumor progression and furthermore bad accessibility to the tumor respectively. In contrast to our current methods, the LB enables us to detect and analyze circulating tumor DNA (ctDNA) in a non-invasive way which solves the problem of inaccessibility to the tumor (see figure 3). Similarly to radiomics we could in combination with next generation sequencing (NGS) detect mutations that are significant for the patient's OS. The promising idea of using liquid biopsies as a biomarker in precicion medicine is further discussed in the following article [17].

LB differs from radiomics in that it can give a better "real time" understanding of the tumor progression due to easier accessibility to the tumor and a lower cost with blood samples over MRI scans. Although the advantages of LB the technique also comes with significant drawbacks. Just like radiomics this is a relatively new emerging fields that have yet to be included independently in clinical use. The biggest drawback of LB alone is the risk of false-negative results and some problems with detection of ctDNA in certain tumor types. Overall, an improvement of analytical method is necessary to increase LBs sensitivity [18].
Despite the independent drawbacks of radiomics and LB, a combination of these minimally invasive circulating biomarkers could be more valuable and reliable than the independent use of the two strategies. Furthermore the combination between these techniques opens the possibilities of new synergies. Two synergies we find particulary intresting and relevant for our project is that 1) Radiomics could ameliorate NGS cost-effectiveness by adding imaging features that reflect underlying gene expression patterns and 2) radiomics could be used to refine LB results and furthermore provide full-field analyses of the patients tumor in a real time response[18].

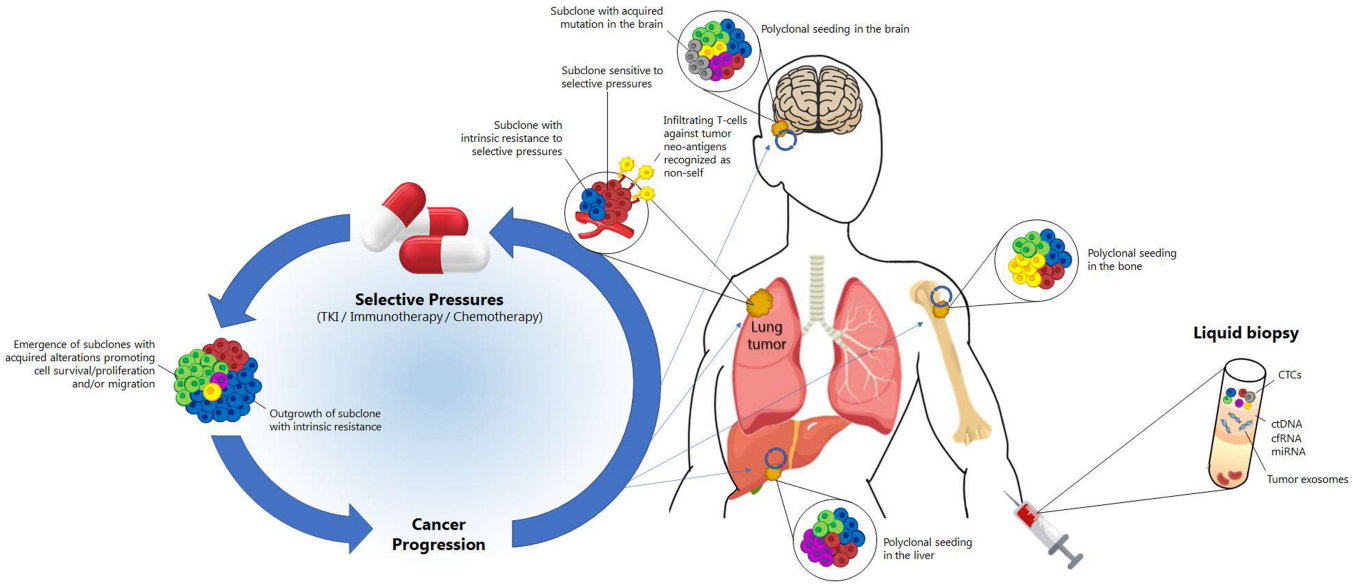

Figure 4: A figure on how the liquid biopsy can assess tumor heterogeneity. This is particularly advantagous for precicion medicine. In the blood sample taken in the right side of the picture, we can isolate tumor markers such as CTCs, ctDNA, cfRNA ang tumor exomes etc. However, our study will only focus on ctDNA[18]

.

1.4 Evaluation

Before the model is used in a clinical setting, it needs to be tested and evaluated thoroughly. Our model will be trained on synthetic data. How we generate the synthetic data and why we use synthetic data is described in section 2.1.1. We will use 5-fold cross validation in the process of training our model. In 5-fold cross validation the synthetic data set is split into five subsets and the model is trained five times. One of the sets are held out and serves as a validation set. The model is trained on the other four subsets. In this way every entry in the synthetic data set will appear in the training set four times and once in the test set and this reduces bias. After the model is trained, we will evaluate how well the model performs by calculating the Root Mean Squared Error (RMSE) between predictions and the original data. To avoid our model becoming a black box where we are unable to explain its predictions, we will use some tools from the field of Explainable AI. This includes looking into feature importance. What features are most important for the predictions of the machine? In this step of the machine learning project life cycle it is important to collaborate with domain experts. If the feature diabetes for instance, proves to be the most important feature for the model, it is necessary to ask a doctor if there really is a causality between diabetes and OS for GBM-patients. If the doctor disagree strongly with the model, we should evaluate our model. Maybe our data set is skewed with a higher proportion of diabetes patients then non-diabetes patients. This is a data set that does not represent the real world and therefore our model makes bad predictions on real data.

References

[1] Ostrom QT, Cioffi G, Gittleman H, Patil N, Waite K, Kruchko C, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2012–2016. Neuro-oncology. 2019;21(Supplement_5):v1-v100.

[2] Patterson JD, Wongsurawat T, Rodriguez A. A glioblastoma genomics primer for clinicians. Medical research archives. 2020;8(2).

[3] American-Brain-Tumor-Association. Types of biomarkers;. https://www.abta.org/wp-content/uploads/2018/03/types-of-biomarkers.pdf.

[4] Louis DN, Perry A, Reifenberger G, Von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. Acta neuropathologica. 2016;131(6):803-20.

[5] Nadeem MW, Ghamdi MAA, Hussain M, Khan MA, Khan KM, Almotiri SH, et al. Brain tumor analysis empowered with deep learning: A review, taxonomy, and future challenges. Brain sciences. 2020;10(2):118.

[6] Chaddad A, Kucharczyk MJ, Daniel P, Sabri S, Jean-Claude BJ, Niazi T, et al. Radiomics in glioblastoma: current status and challenges facing clinical implementation. Frontiers in oncology. 2019;9:374.

[7] Sterckx W, Coolbrandt A, de Casterlé BD, Van den Heede K, Decruyenaere M, Borgenon S, et al. The impact of a high-grade glioma on everyday life: a systematic review from the patient's and caregiver's perspective. European Journal of Oncology Nursing. 2013;17(1):107-17.

[8] Cheon S, Agarwal A, Popovic M, Milakovic M, Lam M, Fu W, et al. The accuracy of clinicians' predictions of survival in advanced cancer: a review. Ann Palliat Med. 2016;5(1):22-9.

[9] Park JK, Hodges T, Arko L, Shen M, Iacono DD, McNabb A, et al. Scale to predict survival after surgery for recurrent glioblastoma multiforme. Journal of clinical oncology. 2010;28(24):3838.

[10] Burger PC, Green SB. Patient age, histologic features, and length of survival in patients with glioblastoma multiforme. Cancer. 1987;59(9):1617-25.

[11] Yang W, Warrington NM, Taylor SJ, Whitmire P, Carrasco E, Singleton KW, et al. Sex differences in GBM revealed by analysis of patient imaging, transcriptome, and survival data. Science translational medicine. 2019;11(473).

[12] Lacroix M, Abi-Said D, Fourney DR, Gokaslan ZL, Shi W, DeMonte F, et al. A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival. Journal of neurosurgery. 2001;95(2):190-8.

[13] Travers S, Litofsky NS. Daily Lifestyle Modifications to Improve Quality of Life and Survival in Glioblastoma: A Review. Brain Sciences. 2021;11(5):533.

[14] Nehme A. Understanding Convolutional Neural Networks; 2018. https://towardsdatascience.com/understanding-convolutional-neural-networks-221930904a8e.

[15] Saha S. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way; 2018. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.

[16] van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging—"How-to" guide and critical reflection. Insights into Imaging. 2020;11(1):1-16.

[17] Bark JM, Kulasinghe A, Chua B, Day BW, Punyadeera C. Circulating biomarkers in patients with glioblastoma. British journal of cancer. 2020;122(3):295-305.

[18] Cucchiara F, Petrini I, Romei C, Crucitta S, Lucchesi M, Valleggi S, et al. Combining liquid biopsy and radiomics for personalized treatment of lung cancer patients. State of the art and new perspectives. Pharmacological Research. 2021:105643.

2 Data management plan and ethical considerations

2.1 Description of generated data and code

The original data we will use in this study will be obtained from Haukeland University Hospital in Bergen, Norway. Patients with GBM with age above 18 years are included in the data-set. The features of the data set are presented in table 1. The data will be stored in a separate secure location and will be linked via an identifier. Any data transfer will be encrypted, the storage devices will use biometric security features, and the backups will be held securely. The original data will be deleted after testing the model. The synthetic data produced by the GAN will be formatted similarly to the original data and stored in three different formats. MRI images will be saved as NIfTI files (.nii.gz), Genomic data will be stored in BAM files, and clinical data will be stored in one CSV file. The images, genomic data and clinical date belonging to the same subject will share the same identifier. Outcome will be overall survival. The data set from Haukeland is at the size of 250 GBM-patients.

Feature	
Age	Numerical
BMI	Numerical
Sex	Boolean
Married	Boolean
Diabetes	Boolean
Neurodegenerative disease	Boolean
Severe fragility	Boolean
Seizure	Boolean
Headache	Boolean
IDH-1 wild type	Boolean
MGMT	Boolean
1p19q codedeletion	Boolean
Histone H3-K27M mutation	Boolean

Table 1: The features we use to predict OS for GBM patients. Histological tumor features as size, shape and location are also included (not shown in the table).

2.1.1 Generating data

In order to create a good model, we will need large amounts of data. One option could be to contact different hospitals around the world and get MRI scans with demographics and other clinical data. That will not only be time consuming, but can also arise problems with different parts of data not corresponding. Our solution to this is to get data from one hospital and use that to create synthetic data.

To create the synthetic data, we will be using Generative Adversarial Networks (GANs). In general, GAN consists of two parts, the generator and the discriminator. The generator model will be trained to generate new images and the discriminator will try to classify the images as either real (from the real data) or fake (from the generated images). This process goes on until the generator creates images that look real enough that the discriminator is fooled [1].

This means generating fake tumor pictures in artificial MRI-pictures and fake patients. The problem with GAN is that it generates data with low variance. Our solution to this is to accumulate all the MRI scans we can get, including scans without GBM, in order to train the model to produce new images. However, based on the paper written by [2], they were able to get good results on creating synthesized PET scans only from

411 images.

With large amount of synthetic data, we will be able to make a deep learning model that surpasses the earlier models made for prediction OS for GBM-patients. We expect to generate a data set of 100 000 patients. Our model will with higher accuracy tell which genes the patient have and therefore provide precision medicine.

2.2 Sharing of data and code

Access to the original data is restricted and only possible through requesting it from the hospital. The synthetic data will be publicly available on figshare.com. The code used to build the model will be freely available on the associated page at github.com.

2.3 Ethical considerations

Obtaining the original data will require ethical approval from the hospital's ethics committee and the Western Norway Committee for Medical and Health Research Ethics. Informed consent must be obtained from patients. The study will be conducted in accordance with the declaration of Helsinki (2000) and the four ethical principles established by the European high-level expert group on artificial intelligence

References

[1] Brownlee, Jason. Generative adversarial networks with python: deep learning generative models for image synthesis and image translation. Machine Learning Mastery.2019

[2] Islam, Jyoti and Zhang, Yanqing, GAN-based synthetic brain PET image generation, Springer, 2020