Precision medicine and quantitative imaging in glioblastoma

ELMED219 (Artificial Intelligence and Computational Medicine), 4-29 January 2021

https://github.com/MMIV-ML/ELMED219-2021

Team #5

1 Research plan

1.1 A brief background to the field

Glioblastoma (GBM) is a tumour of the central nervous system (CNS). GBM, which the WHO classifies as a grade IV glioma [5], is the most malignant form of glioma. From the diagnosis of GBM, the median survival is just 15 months, even with treatment including surgical resection, radiotherapy and chemotherapy. [8, p. 881].

All gliomas are tumours of the brain parenchyma priorly classified as astrocytomas, oligodendrogliomas, and ependymomas based on their morphological resemblance to various types of glial cells. Emerging genetic information has shown that gliomas are a molecularly distinct family of neoplastic lesions, independent of histological patterns. This is reflected in the WHO's updated (2016) categorisation of CNS tumours, which incorporates molecular parameters plus histology. Notably, in case of discordant results, genotype trumps histological phenotype [10]. Thus, the distinction between CNS tumours now hinges on biomarkers, particularly isocitrate dehydrogenase (IDH) mutations and 1p19q co-deletion status [5].

Clinical trials have yet to reveal an effective therapy for most brain tumours, and GBM is no exception. Surgical resection is a difficult process with significant risk to the patient, and since the tumours are located behind the blood-brain-barrier (BBB), they are less exposed to systemic chemotherapy. Finally, "(...) unique developmental, genetic, epigenetic and microenvironmental features of the brain frequently render these cancers resistant to conventional and novel treatments alike" [1].

MRI, magnetic resonance imaging, is used for imaging of GBM and diagnostics. CT, computer tomography, can also be used, however the MRI works better for soft tissue. MRI images are T1-weighted and T2-weighted images, often referred to as T1 and T2 images. T1 images highlight primary fat tissue, while T2 highlights both fat and water. As for the GBM to be pictured, the T2 imaging is preferred [22]. T1 images are useful for detecting hemorrhage of the biopsy. T2 will provide information about the soft tissue resolution for tumor visualization.

In recent years, however, artificial intelligence (AI), particularly deep learning (DL) (advanced machine learning), has developed in leaps and bounds, and there is hope that DL could aid in diagnosis and treatment of patients diagnosed with GBM.

1.2 Objectives and expected impact

The objective of this project proposal is to review the possibility that DL may be used to create a more personalized treatment of GBM patients, by feeding the DL with imaging-derived biomarkers, patient data,

	IDH-wildtype glioblastoma	IDH-mutant glioblastoma	References
Synonym	Primary glioblastoma, IDH-wildtype	Secondary glioblastoma, IDH-mutant	{1830}
Precursor lesion	Not identifiable; develops de novo	Diffuse astrocytoma Anaplastic astrocytoma	{1827}
Proportion of glioblastomas	~90%	~10%	{1797}
Median age at diagnosis	~62 years	~44 years	{214,1078,1797, 2103}
Male-to-female ratio	1.42:1	1.05:1	{214,1417,1797}
Mean length of clinical history	4 months	15 months	{1797}
Median overall survival			
Surgery + radiotherapy	9.9 months	24 months	{1797}
Surgery + radiotherapy + chemotherapy	15 months	31 months	{2810}
Location	Supratentorial	Preferentially frontal	{1417}
Necrosis	Extensive	Limited	{1417}
TERT promoter mutations	72%	26%	{1801,1830}
TP53 mutations	27%	81%	{1797}
ATRX mutations	Exceptional	71%	{1519}
EGFR amplification	35%	Exceptional	{1797}
PTEN mutations	24%	Exceptional	{1797}

Figure 1: Key characteristics of IDH-wildtype and IDH-mutant glioblastomas [12].

treatment and outcomes from earlier cases, and thus, helping clinicians achieve improved survival for their patients.

GBMs consist of different genetic subpopulations, and exhibit great heterogeneity both spatially and temporally. This allows the tumor to adapt to environmental forces. This explains why GBM patients tend to respond poorly to treatment. However, a tumour's genomic alterations develop distinct radiographic phenotypes [7]. This is where MRI plays a key role in characterising molecular signatures of GBM, based on regional heterogeneity and phenotypic presentation of the tumour, in a new research field called radiogenomics [7]. Radiogenomics is advantageous, as it is a noninvasive and global assessment of the tumour and its response to therapies.

The expected impact is to discover interrelationships between tumour biomarkers (as monitored by radiogenomic imaging), in the form of pathological genetic alterations, and tumour therapies currently in use, i.e. radiotherapy, surgery and chemotherapy as temozolomide [23]. As suggested by Stead and Verhaak [23], one way forward could be "larger longitudinal molecular profiling studies" using multi-tumour analysis. We propose a supplemental role for DL to explore possible interrelationships, based on information from such studies, to uncover potential combinations of variables conducive to increased survival. The impact is expected to be increased survival, as DL informs a more personalised, "tailored" use of current therapies.

In short, the objective of this project proposal is to increase survival by avoiding merely "pruning" tumours, which allows for recurrence. The expected impact is for DL to allow therapies to nip tumours in the proverbial bud. It is our hope that DL, employing information from radiogenomic imaging and other clinical data, can help clinicians discern the optimal treatment strategy for each patient, thus eradicating all cells bearing detrimental genetic signatures.

1.3 Material and methods

After the patient is diagnosed or treated, we will have MRI images, biopsy and molecular markers for each GBM case. Our goal is to achieve improved 5-year-survival rates for the patients with GBM, and therefore we want to use DL to give us a better treatment plan. If the DL can analyze MRI images, biopsy and biomarkers, and tell how long a patient with those values lived, then the DL may create a better treatment plan for the values we are feeding the DL with. The data DL is fed with, is for patients over 20 years diagnosed with GBM.

1.3.1 Deep learning

Deep learning (DL) is a type of AI that, as opposed to traditional machine learning, can learn from data without hard-coded rules or human knowledge. This means that by inputting raw data such as images, text, audio and/or video, DL can automatically learn and grasp aspects from it. By increasing the amount of data the DL has access to, one can improve the DL's accuracy. Today, DL is used throughout a variety of apps and software including computer vision, conversational AI and recommendation systems. The latter use images, texts and other inputs to offer meaningful advice on for example how to finish a sentence or, in our case, what treatment to give a patient [20]. In our project, we will use supervised deep learning to achieve a treatment as an output, by comparing the current input (MRI images, biomarkers and other data) to previous input-output pairs.

1.3.2 Segmentation and classification

Segmentation is to sort the data into groups. A well defined segment includes data with similar features, while the features within different segments differ from each other. One of the biggest challenges by segmentation is to decide how many segments one should divide the data into. Segmentation is an important part of MRI image processing, when finalising the image analysis.

Classification is a form of under-categorization within each segment, which divides the data into even smaller groups called classes. To do this, it is necessary to use machine learning algorithms that are able to understand which classes to annotate the different examples to. An example would be classifying email as spam or not spam [2].

Classes are listed in 2.1. The software will use binary and multi-class classification. These features will be independent, and not have an effect on each other.

To help the program understand the MRI images, the data needs to be pre-processed before it can be used as an input to the software. It is especially important to decide on a number of pixels and cells for all images, making it possible for the computer to compare them. We start by inputting the MRI images, and then applying various filters and adjustments including image contouring, cropping, resizing and normalization. We continue by making a histogram of oriented gradients (HOG) to get the pre-processed image as an output. Python should be able to read our histogram of oriented gradients [9] when it's processed. The goal is to turn the MRI images into meaningful regions for the software to read. Picture quality will most likely vary due to patient treatment in different facilities, therefore input data will have to be pre-processed carefully. [29]. Convolutional neural network (CNN) will be one way to make this happen. "Each neuron in a neural network computes an output value by applying a specific function to the input values coming from the receptive field in the previous layer. The function that is applied to the input values is determined by a vector of weights and a bias (typically real numbers). Learning consists of iteratively adjusting these biases and weights".

In some cases the information and data available will vary. This may cause an imbalance in the dataset. However, we hope to train the software to overlook the skewed amount of data, and be able to do predictions even though the patient presents with "missing" parameters.

1.3.3 Treatment stratification

Stratification of clinical studies is to partition test subjects and results with another factor than the given treatment. This can be done based on sex, age or other demographical factors [31]. Considering some patients do not have the same healthcare options, it will be reasonable to think treatment facility and hospital will be some of the influential factors. Due to a shortage of confounding variables available in the dataset this might be a challenge to determine. Considering the number of patients, difference in healthcare given, and demographic factors, the projection of the discovered treatment stratification might include some errors.

1.3.4 Prediction

After the model and software are completed, training and testing will be necessary to achieve the appropriate results. We can start this process by doing a train-test split on our data, which divides it into a training set and a testing set. We will then train the software on the training data, before we test it on the test data. After this is done successfully, it is important to test the software on real world cases as well, where we already know the outcomes, and evaluate how precise the program's ability is to predict the patients' treatment and outcome. Our goal is a software which chooses the right medication and treatment in at least 90% of the cases. The software will improve the more cases it has seen. As mentioned in 1.3.2, some data available might not be of the quality one desires. In time, this should improve, and so will the software. The software output and result will determine the preferred treatment for a patient with GBM, given the wanted outcome is to increase days to death. It will be up to the treatment team to interpret the software predictions. In some cases, changing treatment facilities might be suggested. However, this might not be an option. Hopefully the software will find correlations between patient data, and optimise treatment.

1.3.5 Limitations of Deep Learning

The software is not able to make predictions creatively and with regard to clinical concerns, like a human doctor. This could, arguably, limit the applicability of DL predictions. An example of this limitation, termed underspecification, is discussed below in the ethics paragraph.

Old data cannot be used in the same way as new data. By comparing non-recent computer images to the ones we have today, it's obvious there has been a clear development in imaging. So, quality can vary between new and old images, especially with a HOG in mind, rendering the HOG imprecise. Using a high number of pixels can make the old images inadequate, making the software fail to differentiate between pathological and healthy tissue. Even though this might not be DL's fault, despite having high amounts of data as input, it's still a limitation for the software.

We should have follow-up images, and favourably several MRI images during the treatment, for the software to better understand how the tumor is evolving. We then imagine that the software could predict which functions and abilities will be affected by the tumor, and thus decide on which treatment can best prevent the loss of functions.

1.4 Evaluation

Our goal is to improve general GBM survival. Given this goal, we believe that the project outline could improve diagnosis and treatment of this patient group. Our ambition is tempered by data availability and DL functionality and applicability. As patients continue to pass away or survive, we cannot add or update how long they have stayed alive, or how long they will continue to do so. Ultimately, the training data is a snapshot of a moment in time in a patient history. Therefore, we conclude that the utility of DL in our project outline is as a diagnostic aide to human doctors.

2 Data management plan and ethical considerations

2.1 Description of generated data and code

We will be using data of patients diagnosed at the age of 20 and older. The data is collected from the TCGA-GBM data collection [18], from the GDC Data Portal from the National Cancer Institute [19] and Cancer Imaging Archive databases [25]. We will also use necessary data from electronic patient journals and HelseNorge to supply the biomarkers and cases to further train and test the DL. For that age group, the data collection from TCGA-GBM, contains 590 cases, where 486 are dead, 101 are still alive and 3 are not reported. 20 patients died within a month after being diagnosed and 230 patients died within a year. 469 out of 590 patients were reported dead within five years after diagnosis, which corresponds to 79,5%. The data we use is from 2014 to 2020.

Data used for DL are described in table 1. Parameters are shown on the left side, and levels are shown to the right. In addition to the listed parameters, all available MRI images of the tumors and detailed description of the conducted treatment, will be included in the machine learning program. The MRI-pictures are filtered, as described in 1.3.3. MGMT-methylation is, as of today, an important test to decide on the prognosis and which treatment to give the patients. We hope that the status of a possible IDH-mutation will also have an impact on the treatment or estimated prognosis. To test for MGMT-methylation and IDH-mutation, tumor tissue is necessary [6][27]. MGMT-methylation, IDH-mutation, 1p19q co-delection status and MRI images are all used to determine diagnosis. These variables of a new patient are to be compared to the same variables of the patients from the data collection, their treatment and the result of the treatment in terms of survival. Their survival is stored as the number of days they survived after receiving the diagnosis. The treatment can be done in many ways, and can consist of surgery, chemotherapy and radiation. Which of the treatment methods that has been used, is important to use as input for DL . If chemotherapy is used, which type, dose and duration must also be included. Also if it is used before, during and/or after any surgery. As for radiation treatment, GBM is treated with IMRT [16], photon, x-ray [24], proton therapy and gamma radiation from gamma knife. If a patient is treated with a form of radiation, then what kind of radiation, energy, fractions and dose is also included. Gender and age are also parameters. This is to see if there is any correlation between gender, age and the most appropriate treatment.

Parameter	**Level**
MGMT-metylert tumor	True/False
Isocitrate dehydrogenase (IDH) mutation	True/False
1p19q co-deletion status	True/False
Age	20-29, 30-39, 40-49, 50-59, 60-69, 70$\leq$
Gender	Male/Female
Survival after diagnosis	Number of days

Table 1: List of parameters

2.2 Sharing of data and code

As mentioned in 2.1, for this project we plan on using already publicly available data, from the GDC Data Portal from the National Cancer Institute [19], Cancer Imaging Archive databases [25] and electronic patient journals (HelseNorge) to train and test the software. Sharing the code will be done open-source through GitHub using the MIT license and by providing a Jupyter notebook to better explain the code. By providing public access through GitHub, others will also be able to edit the code and create forks based on it.

2.3 Ethical considerations

In applying Deep Learning (DL) to combat glioblastoma (GBM), one must not forget ethics. The Hippocratic Oath forms the basis of modern medical ethics, which in its abbreviated form reads: "primum non nocere, secundum cavere, tertium sanare" - first, do no harm; second, be "cautious"/prevent disease; third, cure [15][13]. In the context of AI systems, we also apply the Ethics Guidelines for Trustworthy Artificial Intelligence (AI): respect for human autonomy, prevention of harm, fairness, and explicability [4].

DL is a promising new tool for assessing GBM. DL does, however, have limitations, which stem from theoretical and observed limitations inherent in DL as a method (e.g. a tendency to reinforce inherent biases from the training data); a lack of what is deemed sufficient amounts of training data ("big data"); and legal responsibility with a view to treatment: while entrepreneurs may think big and take risks, it remains the doctors' job to protect their patients [18].

An inherent DL limitation is a newly described phenomenon called underspecification, which occurs when DL is applied to complex real-world applications. In general, the solution to a problem is underspecified if there are many distinct solutions (pairs of values) that solve a problem equivalently. However, these pairs can lead to dramatically different predictions. "The problem arises because the machine learning process has no way to properly choose between these pairs. Indeed, (...) the parameters the machine chooses can depend on entirely arbitrary decisions in the way the model is set up" [3].

This places limitations on the credibility of machine learning predictions and may force some adjustments in its applications. This is a particular concern in medical imaging [3].

Training data is necessary input experience for DL, but this input contains biases from human-made systems. Patient data in the TCGA-GBM data collection is collected from patients all over the world [26]. However, there is a higher availability of data from Caucasian patients relative to other ethnic groups [17]. Considering sex-dependent patient variation in GBM genomics, immune response and imaging characteristics, DL could thus be missing out on variations between ethnic groups, rendering its predictions less valuable for non-Caucasian patients [7, p. 5]. The ethnic imbalance in the dataset could, however, simply reflect the fact that GBM affects more Caucasians than other ethnic groups [21]. In the end, the lower prevalence of GBM in certain ethnic groups may lead to poorer machine learning predictions for these patients.

Amongst entrepreneurs and clinicians, self-interest in the form of employment, fear of change or competition over research grants is another ethical variable [1]. Designing research collaboration with this in mind can help skew results towards innovation and better patient care, rather than the opposite.

Finally, there are concerns surrounding the use and handling of patient data as input. With the development of DL, there is a need for more training data, in quantities that have given rise to the concept of "big data". In the TCGA-GBM data set, this is addressed in part by way of anonymisation. Patients should also consent to sharing of such data, and be confident that their data is handled safely and confidentially. There have been some problems in this area, concerning confidential and safe data handling [28].

References

[1] Aldape, K, Brindle, KM, Chesler, L et al. Challenges to curing primary brain tumours. Nat Rev Clin Oncol. 2019;16: 509–520. https://doi.org/10.1038/s41571-019-0177-5 Available at: https://www.nature.com/articles/s41571-019-0177-5

[2] Cs.Cornell. Lecture 1: supervised learning [Online]. [03.11.17; 25.01.21]. Available at: http://www.cs.cornell.edu/courses/cs4780/2018fa/lectures/lecturenote01_MLsetup.html

[3] Discover Magazine. Google Reveals Major Hidden Weakness In Machine Learning [online]. Waukesha: Kalmbach Media; 2021. [Updated 30 Nov 2020; accessed 26 Jan 2021]. Available at: https://www.discovermagazine.com/technology/google-reveals-major-hidden-weakness-in-machine-learning

[4] European Commission. Ethics Guidelines for Trustworthy AI [online]. Brussels: European Commission; 2021. [Accessed 26 Jan 2021]. Available at: https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines

[5] Gaillard, F. et al. Diffuse glioma [online]. Radiopaedia; 2021. [Accessed 26 Jan 2021] Available at: https://radiopaedia.org/articles/diffuse-glioma

[6] HistoGeneX. MGMT Testing (Glioblastoma). [online]. Antwerp: HistoGeneX; 2019. [Updated 23 May 2019; accessed 26 Jan 2021]. Available at: https://www.histogenex.com/mgmt-testing-glioblastoma

[7] Kazerooni, AK, Bakas, S, Rad, HS, Davatzikos, C. Imaging Signatures of Glioblastoma Molecular Characteristics: A Radiogenomics Review. Journal of Magnetic Resonance Imaging (JMRI). 2019;52(1): 1-16. DOI: 10.1002/jmri.26907 Available at: https://qmisg.com/wp-content/uploads/Imaging-Signatures.pdf

[8] Kumar V, Abbas AK, Aster JC. Robbins Basic Pathology. 10th ed. Pennsylvania: Elsevier; 2016. 935 pp.

[9] Learn OpenCV. Histogram of Oriented Gradients [Internett]. Satya Mallick; [06.12.16; 21.01.21]. Available at: https://learnopencv.com/histogram-of-oriented-gradients/)

[10] Louis, DN, Perry, A, Reifenberger, G et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol. 2016; 131: 803–820. [Accessed 26 Jan 2021]. Available at: https://link.springer.com/article/10.1007/s00401-016-1545-1

[11] Louis, DN, Perry, A, Reifenberger, G et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary: Table 1. Acta Neuropathol. 2016; 131: 803–820. [Accessed 26 Jan 2021]. Available at: https://link.springer.com/article/10.1007/s00401-016-1545-1/tables/1

[12] Louis, DN, Perry, A, Reifenberger, G et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary: Table 4. Acta Neuropathol. 2016; 131: 803–820. [Accessed 26 Jan 2021]. Available at: https://link.springer.com/article/10.1007/s00401-016-1545-1/tables/4

[13] Mannheimer, GL. Primum non nocere [online]. San Francisco: Medium; 2019. [Updated 2021; accessed 26 Jan 2021]. Available at: https://medium.com/hola-glovo/primum-non-nocere-2c2d71d46c9c

[14] MD Anderson Cancer Center. Proton therapy treatment for glioblastoma [online]. Houston: MD Anderson Cancer Center; 2021. [Accessed 26 Jan 2021]. Available at: https://www.mdanderson.org/publications/cancer-newsline/proton-therapy-for-glioblastoma.h35-1589046.html

[15] MedicineNet. Medical Definition of Hippocratic Oath [online]. New York City: WebMD; 2021. [Accessed 26 Jan 2021]. Available at: https://www.medicinenet.com/hippocratic_oath/definition.htm

[16] Moffitt Cancer Center. Glioblastoma Radiation [online]. Tampa: Moffitt Cancer Center; 2018. [Accessed 26 Jan 2021]. Available at: https://moffitt.org/cancers/glioblastoma/treatment/radiation/

[17] National Cancer Institute: GDC Data Portal. Exploration: TCGA: TCGA-GBM [online]. Washington, D.C.: U.S. Department of Health and Human Services; 2020.[Updated 29 Oct 2021; accessed 26 Jan 2021]. Available at: https://portal.gdc.cancer.gov/repository?facetTab=cases&filters=%7B%22op%22%3A%22and%22%2C%22content%22%3A%5B%7B%22content%22%3A%7B%22field%22%3A%22cases.demographic.days_to_death%22%2C%22value%22%3A%5B%22365%22%5D%7D%2C%22op%22%3A%22%3E%3D%22%7D%2C%7B%22content%22%3A%7B%22field%22%3A%22cases.diagnoses.age_at_diagnosis%22%2C%22value%22%3A%5B7305%5D%7D%2C%22op%22%3A%22%3E%3D%22%7D%2C%7B%22op%22%3A%22in%22%2C%22content%22%3A%7B%22field%22%3A%22cases.project.project_id%22%2C%22value%22%3A%5B%22TCGA-GBM%22%5D%7D%7D%5D%7D&searchTableTab=cases

[18] National Cancer Institute: GDC Data Portal. Repository: TCGA-GBM [online]. Washington, D.C.: U.S. Department of Health and Human Services; 2020.[Updated 29 Oct 2021; accessed 26 Jan 2021]. Available at: https://portal.gdc.cancer.gov/repository?facetTab=cases&filters=%7B%22op%22%3A%22and%22%2C%22content%22%3A%5B%7B%22content%22%3A%7B%22field%22%3A%22cases.demographic.days_to_death%22%2C%22value%22%3A%5B%22365%22%5D%7D%2C%22op%22%3A%22%3E%3D%22%7D%2C%7B%22content%22%3A%7B%22field%22%3A%22cases.diagnoses.age_at_diagnosis%22%2C%22value%22%3A%5B7305%5D%7D%2C%22op%22%3A%22%3E%3D%22%7D%2C%7B%22op%22%3A%22in%22%2C%22content%22%3A%7B%22field%22%3A%22cases.project.project_id%22%2C%22value%22%3A%5B%22TCGA-GBM%22%5D%7D%7D%5D%7D&searchTableTab=cases

[19] National Cancer Institute: GDC Data Portal. TCGA-GBM [online]. Washington, D.C.: U.S. Department of Health and Human Services; 2020.[Updated 29 Oct 2021; accessed 26 Jan 2021]. Available at: https://portal.gdc.cancer.gov/projects/TCGA-GBM

[20] NVIDIA Developer. Deep Learning [online]. Santa Clara: NVIDIA Corporation; 2021. [Accessed 26 Jan 2021]. Available at: https://developer.nvidia.com/deep-learning

[21] Ostrom, QT, Cote, DJ, Ascha, M, Kruchko, C, Barnholtz-Sloan, JS. Adult Glioma Incidence and Survival by Race or Ethnicity in the United States From 2000 to 2014. JAMA Oncol. 1 Sep 2018; 4(9): 1254-1262. doi: 10.1001/jamaoncol.2018.1789. Available at: https://pubmed.ncbi.nlm.nih.gov/29931168/

[22] Radiology masterclass. MRI interpretation: T1 v T2 images [online]. Salisbury: Department of Radiology, New Hall Hospital; 2021. [Accessed 26 Jan 2021]. Available at: https://www.radiologymasterclass.co.uk/tutorials/mri/t1_and_t2_images

[23] Stead, LF, Verhaak, RGW. Doomed from the TERT? A Two-Stage Model of Tumorigenesis in IDH-Wild-Type Glioblastoma. Cancer Cell. 2019;35(4): 542-544. DOI: https://doi.org/10.1016/j.ccell.2019.03.009 https://www.cell.com/cancer-cell/fulltext/S1535-6108(19)30154-0

[24] Szabo, L. A Reality Check On Artificial Intelligence: Are Health Care Claims Overblown? [online]. San Francisco: Kaiser Family Foundation; 30 Dec 2019. [Accessed 26 Jan 2021]. Available at: https://khn.org/news/a-reality-check-on-artificial-intelligence-are-health-care-claims-overblown/

[25] The Cancer Imaging Archive (TCIA). Browse Data Collections [online]. Little Rock: University of Arkansas for Medical Sciences (UAMS); 2021. [Accessed 26 Jan 2021]. Available at: https://www.cancerimagingarchive.net/collections/

[26] The Cancer Imaging Archive (TCIA). TCGA-GBM [online]. Little Rock: University of Arkansas for Medical Sciences (UAMS); 2020. [Updated 16 Dec 2020; accessed 26 Jan 2021]. Available at: https://wiki.cancerimagingarchive.net/display/Public/TCGA-GBM

[27] UCSF Health Center for Clinical Genetics and Genomics. IDH1 Mutation Testing. [online]. San Francisco: University of California San Francisco; 2021. [Accessed 26 Jan 2021]. Available at: https://genomics.ucsf.edu/content/idh1-mutation-testing

[28] Vincent, J. Google's DeepMind made 'inexcusable' errors handling UK health data, says report [online]. New York: Vox Media; 16 March 2017. [Accessed 27 Jan 2021]. Available at: https://www.theverge.com/2017/3/16/14932764/deepmind-google-uk-nhs-health-data-analysis

[29] Wikipedia. Convolutional neural network [Online]. [17.01.21; 25.01.21]. Available at: https://en.wikipedia.org/wiki/Convolutional_neural_network

[30] Wikipedia. Hippocratic Oath [online]. San Francisco: Wikimedia Foundation; 2021. [Updated 25 Jan 2021; accessed 26 Jan 2021]. Available at: https://en.wikipedia.org/wiki/Hippocratic_Oath

[31] Wikipedia. Stratification (clinical trials) [Online]. [30.12.20;23.01.21]. Available at: https://en.wikipedia.org/wiki/Stratification_(clinical_trials)#:~:text=Stratification%20of%20clinical%20trials%20is,age%2C%20or%20other%20demographic%20factors