Imaging and Resection of Glioblastoma

in light of molecular markers

Fatih Incekara

Imaging and Resection of Glioblastoma

in light of molecular markers

Fatih Incekara

ISBN: 978-94-6423-080-2

Design by: Fatih Incekara and Dennis Hendriks / ProefschriftMaken.nl
Cover and chapter illustrations used with permission of Kenhub.com
Printed by: ProefschriftMaken.nl

Imaging and Resection of Glioblastoma

in light of molecular markers

Beeldvorming en Resectie van Glioblastoom

in het licht van moleculaire markers

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the Rector Magnificus
Prof.dr. F.A. van der Duijn Schouten

and in accordance with the decision of the Doctorate Board.

The public defense shall be held on
Wednesday 10th of March 2021
at 13.00 hours

by

Fatih Incekara
born in Rotterdam

Erasmus University Rotterdam

Promotors
Prof. dr. M. Smits
Prof. dr. M.J. van den Bent

Other members
Prof. dr. P.C. de Witt Hamer
Prof. dr. M.W. Vernooij
Prof. dr. M.J.B. Taphoorn

Copromotor
Dr. A.J.P.E. Vincent

to Sara
to Merve
to my parents

dare to know
Immanuel Kant

know thyself
Socrates

he who knows himself knows his Lord
Prophet Mohammed
peace be upon him

Contents

List of publications

1. Intraoperative Ultrasound Guidance and Extent of Resection in Glioblastoma Surgery: a Randomised, Controlled Trial
   ***Fatih Incekara**, Marion Smits, Linda Dirven, Eelke Bos, Rutger Balvers, Iain Haitsma, Joost Schouten, Arnaud J.P.E. Vincent.*
   Submitted

2. The association between the extent of glioblastoma resection and survival in light of MGMT promoter methylation in 326 patients with newly diagnosed IDH wildtype glioblastoma.
   ***Fatih Incekara**, Marion Smits, Sebastian R. van der Voort, Hendrik Jan Dubbink, Peggy N. Atmodimedjo, Johan M. Kros, Arnaud J.P.E. Vincent, Martin van den Bent.*
   Front. Oncol. 2020 Jul 10;10:1087.

3. Topographical Mapping of 436 Newly Diagnosed IDH Wildtype Glioblastoma with vs. without MGMT Promoter Methylation
   Fatih Incekara*, *Sebastian van der Voort*, Hendrikus J. Dubbink, Peggy N. Atmodimedjo, Rishi Nandoe Tewarie, Geert Lycklama, Arnaud J.P.E. Vincent, Johan M. Kros, Stefan Klein, Martin van den Bent, Marion Smits. Front.*
   Oncol. 2020 May 12: 10:596.

4. Predicting the 1p/19q co-deletion status of presumed low grade glioma with an externally validated machine learning algorithm
   Sebastian van der Voort, **Fatih Incekara***, Maarten M.J. Wijnenga, Georgios Kapas, Mayke Gardeniers, Joost W. Schouten, Martijn P.A. Starmans, Rishi Nandoe Tewarie, Geert J. Lycklama, Pim J. French, Hendrikus J. Dubbink, Martin van den Bent, Arnaud J.P.E. Vincent, Wiro J. Niessen, Stefan Klein, Marion Smits.*
   Clin Cancer Res. 2019 Dec 15;25(24):7455-7462.

5. Association between supratotal glioblastoma resection and patient survival: a systematic review and meta-analysis
   ***Fatih Incekara**, Stephan Koene, Arnaud J.P.E. Vincent, Martin van den Bent, Marion Smits.*
   World Neurosurg. 2019 Jul;127:617-624.

6. Letter to the Editor: Supratotal resection of glioblastoma
   ***Fatih Incekara**, Marion Smits, Arnaud J.P.E. Vincent,*
   J Neurosurg. 2019 Aug 16:1-2.

7. Clinical feasibility of a wearable mixed reality device in neurosurgery
***Fatih Incekara**, Marion Smits, Clemens Dirven, Arnaud J.P.E. Vincent.*
World Neurosurg. 2018 Oct;118:e422-e427.

8. Changes in language white matter tract microarchitecture are associated with cognitive deficits in patients with presumed low grade glioma.
***Fatih Incekara**, Djaina Satoer, Evy Visch-Brink, Arnaud J.P.E. Vincent.*
J Neurosurg. 2018 Jun 8:1-9.

9. WHO 2016 subtyping and automated segmentation of glioma using multi-task deep learning.
Sebastian R. van der Voort, **Fatih Incekara*** , Maarten MJ Wijnenga , Georgios Kapsas, Renske Gahrmann , Joost W Schouten , Rishi Nandoe Tewarie, Geert J Lycklama, Philip C De Witt Hamer, Roelant S Eijgelaar, Pim J French, Hendrikus J Dubbink, Arnaud JPE Vincent, Wiro J Niessen, Martin J van der Bent, Marion Smits*, and Stefan Klein*.*
Submitted

10. Development and external validation of a clinical prediction model for survival in glioblastoma patients.
*Hendrik-Jan Mijderwijk, Daan Nieboer, **Fatih Incekara**, Berger K, Daniel Hänggi, Michael S Sabel, Jörg Felsberg, Guido Reifenberger, Martin J van den Bent, Marion Smits, Marie-Therese Forster, Marcel A Kamp.*
Submitted

11. The Value of Pre- and Intraoperative Adjuncts on the Extent of Resection of Hemispheric Low-Grade Gliomas: A Retrospective Analysis.
***Fatih Incekara**, Olutayo Olubiyi, Aysegul Ozdemir, Tomas Lee, Laura Rigolo, Alexandra Golby.*
J Neurol Surg A Cent Eur Neurosurg. 2016 Mar;77(2):79-87.

12. Intraoperative Magnetic Resonance Imaging in Intracranial Glioma Resection: A Single-Center, Retrospective Blinded Volumetric Study.
*Olutayo Olubiyi, Aysegul Ozdemir, **Fatih Incekara**, Yanmei Tie, Parviz Dolati, Liangge Hsu, Sandro Santagata, Zhenrui Chen, Laura Rigolo, Alexandra Golby.*
World Neurosurg. 2015 Aug;84(2):528-36.

* contributed equally

Chapter 1

Introduction

Glioma epidemiology & clinical presentation

Gliomas are brain tumors that arise from glial cells and are molecularly classified following the updated World Health Organization (WHO) 2016 Classification of Tumors of the Central Nervous System. The main types of diffuse, non-circumscript glioma are oligodendroglioma, astrocytoma and glioblastoma. It is estimated that in 2020, over 20,000 patients will be newly diagnosed with some type of glioma in the United States.[1] Glioblastoma account for the majority of these tumors (57.3%) and are the most aggressive type. The age-adjusted incidence rate of glioblastoma is 3.22 per 100,000 population. The median age of patients diagnosed with glioblastoma is 65 years, with highest rates between 75-84 years. Glioblastoma is 1.58 times more common in men than in women. The etiology of glioblastoma is unknown.

The clinical presentation of brain tumors depends on tumor localization and growth rate. Diffuse astrocytoma and oligodendroglioma (low grade glioma) tend to grow more slowly than glioblastoma; they present less commonly with focal neurological deficits and more often with seizures. Glioblastoma patients present in general more often with sub-acute symptoms that progress over days to weeks, which include persistent headache, fatigue, and focal neurological symptoms, such as memory loss, motor, speech or visual deficits, cognitive and personality changes.[2,3] Seizures are less common in glioblastoma than in low grade gliomas. Magnetic resonance imaging (MRI) is needed for the radiological diagnosis of a cerebral mass lesion.

Patients with a glioma as seen on MRI are referred for surgery on a short term with the aim to undergo maximal safe tumor resection to reduce symptoms, increase survival and ultimately to obtain definitive histopathological and molecular diagnosis. If resection is deemed not feasible, a biopsy is required for tissue diagnosis. After surgery, patients are treated with a radiotherapy and/or chemotherapy scheme, which depends on factors such as age, neurological status, extent of tumor resection and molecular classification of the tumor.

Molecular classification

The WHO 2016 classification is predominantly based on molecular characteristics, in particular mutations in the gene encoding for isocitrate dehydrogenase (IDH) 1 and 2 and 1p/19q codeletion.[4,5] Oligodendroglioma are now defined as diffuse glioma with 1p19q codeletion and IDH mutation; astrocytoma are classified according to their IDH mutational status as either IDH mutated (mt) or wildtype (wt). The 2016 WHO currently distinguishes between glioblastoma IDHwt and IDHmt.[4,6]

More recently, low grade astrocytoma IDHwt with certain molecular characteristics (combined gain of chromosome 7 and loss of chromosome 10 and/or EGFR amplification and/or TERT promoter mutations) are called astrocytoma with molecular features of

glioblastoma. Another more recent change is the renaming of glioblastoma IDHmt as grade IV astrocytoma IDHmt.[7,8]

IDH mutations are early mutations affecting codon 132 in 90 percent of all IDH mutations in diffuse glioma. The mutation leads to changes in the enzyme and consequently in increased levels of 2-hydroxygluteratate and decreased levels of α-ketogluterate and NADPH.(9) Due to these alterations, and due the MGMT promotor methylating effect of IDH mutation, these tumors become more sensitive for alkylating chemotherapy and radiotherapy.[6,10,11] In tumors that accumulate IDH mutation, a combined deletion of the short arm of chromosome 1 and the long arm of chromosome 19 may occur as a result of balanced translocation (oligodendroglioma).[12,13] Next to IDH mutation, 1p19q codeletion is also associated with increased sensitivity for alkylating chemotherapy.[14,15] In glioblastoma, another alteration that is associated with improved prognosis is methylation of the promoter region of the gene O^6-methylguanine-DNA methyltransferase (MGMT).[16-18] MGMT is a DNA repair enzyme, which is expressed by the MGMT gene located on chromosome 10q26. Promoter methylation of this gene reduces MGMT protein expression and consequently decreases DNA repair and increases alkylating chemotherapy induced tumor death. Therefore, patients with MGMT methylated glioblastoma are more sensitive to temozolomide than those without MGMT methylated glioblastoma and thus have a better prognosis. MGMT promoter methylation is present in approximately 35-50% of patients with newly diagnosed glioblastoma.[19] IDH mutation, 1p19q codeletion and MGMT promoter methylation are all associated with more favorable prognosis in patients with glioma.[13,15,20-23]

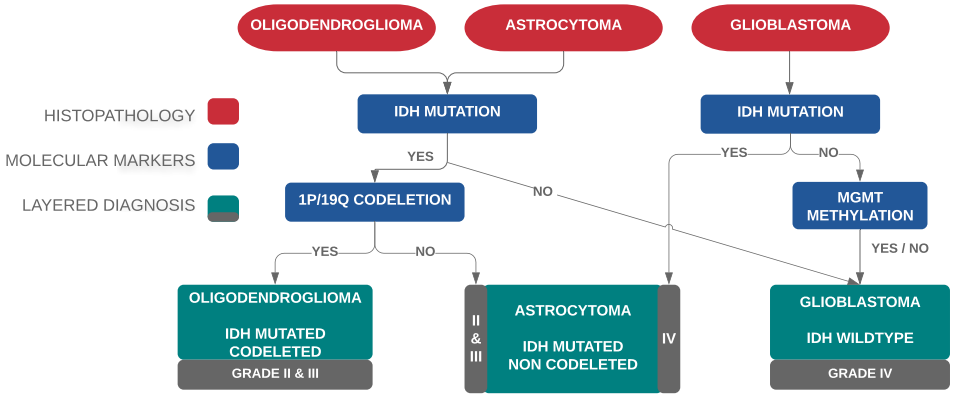


Figure 1. WHO 2016 molecular classification of glioma with cIMPACT-NOW update (brief overview).

Imaging and radiogenomics

Diffuse olidogdendroglioma and astrocytoma are hypointense on T1-weighted MRI scans and hyperintense on T2-weighted MRI scans. Radiogenomics research has the main goal to correlate anatomical and physiological MRI features with molecular subtypes, increasingly with an artificial intelligence approach.[24-31] Studies have indicated that oligodendroglioma with 1p/19q codeletion and IDH mutation are typically located in the frontal lobes with calcification, cortical-subcortical involvement, a heterogeneous appearance on T2-weighted MRI scans with indistinct borders and minimal or patchy contrast enhancement on contrast-enhanced T1-weighted MRI scans. In contrast, astrocytomas with IDH mutation and without 1p/19q codeletion are more often located in the temporal lobe or insular regions. They are homogenous on T2-weighted MRI scans with distinct borders, and they lack calcifications, cortex involvement or contrast-enhancement on contrast-enhanced T1-weighted MRI scans.

Glioblastoma typically appear as a contrast enhancing lesion with central necrosis on post-contrast T1-weighted MRI scans.[29] Glioblastoma infiltrate far beyond the margins of contrast enhancement and together with edema, this infiltration appears as a non-contrast enhanced, hyperintense area on a T2-weighted or T2-FLAIR MRI scan.[32] Advanced and physiological MRI (diffusion weighted imaging and PET-MRI) is shown to be useful to detect glioma infiltration more accurately.[33] There are currently no reliable MRI features that can distinguish MGMT promotor methylated glioblastoma from unmethylated tumors.

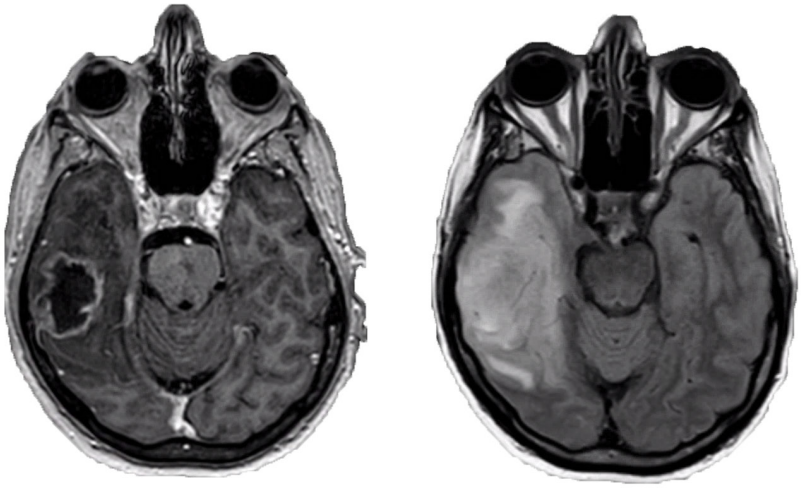


Figure 2. Glioblastoma in the right temporal lobe as seen on post-contrast T1-weighted MRI scan (left) and FLAIR-weighted MRI scan (right).

Extent of resection

In glioblastoma, complete resection of contrast-enhancing tumor on post-contrast T1-weighted MRI has consistently been associated with longer overall survival.[34-39] A systematic review and meta-analysis of 37 articles with over 41.000 glioblastoma patients showed that complete tumor resection decreased the risk of one and two year mortality, when compared to subtotal resection.[34] In addition, more recent studies have shown that resection beyond the borders of contrast enhancement is associated with improved overall survival in patients with glioblastoma.[35, 36, 40-42]. However, complete tumor resection and maximizing resection beyond the borders of contrast enhancement should not be achieved at all cost. Reports on the safety of supratotale resection (beyond the contrast enhancing borders) are still limited in numbers and therefore, needs further investigation.[42] Intraoperative imaging technologies can be used to achieve safe and maximal tumor resection during glioblastoma surgery.[43-45] Intraoperative, real time imaging is needed, since neuronavigation systems are typically based on preoperative MRI scans and due to brain shift, their accuracy in representing the actual situation during surgery decreases. Two randomized controlled trials have shown that 5-aminolevulinic acid and intraoperative MRI guided surgery improves the extent of glioblastoma resection.[43-45] However, an intraoperative MRI system is expensive and prolongs surgery time with approximately one hour.[45] Alternative time- and cost- effective imaging technologies may be useful, such as intraoperative Raman spectroscopy or intraoperative ultrasound guidance.[43,46,47] However, no randomized controlled trial has assessed their value to improve the extent of glioblastoma resection and overall outcome.[43]

Despite improved surgical and imaging techniques, maximization of the extent of resection and the addition of temozolomide to radiotherapy over the past few decades, glioblastoma patients still have a poor prognosis of 15 months (6.8% five-year overall survival rate).[1,48-50] Patients eventually show disease progression and die due to mass effect or extensive brainstem infiltration.[51,52] There is currently no cure for glioblastoma.

Aims and outline of thesis

In this thesis, we assessed the value of glioblastoma imaging and resection in light of molecular markers.

In **Part I** of this thesis, our aim was to predict molecular markers of glioma on preoperative MRI scans. In **Chapter 2** we voxel-wise analyze whether there is a difference in anatomical localization between IDH wildtype glioblastoma with vs. without MGMT promoter methylation. In **Chapter 3 and 4** we predict molecular subtypes of glioma (i.e. 1p/19q codeletion, IDH mutation, MGMT promoter methylation) based on MRI scans using

machine and deep learning algorithms. In **Chapter 5** we evaluate cognitive functions of patients with glioma prior to surgery using white matter fiber tracking.

In **Part II** of this thesis, we assessed the value of image guided glioblastoma resection. In **Chapter 6** we present the results of the ultrasound trial, which is a randomized controlled trial that assesses the value of intraoperative ultrasound guided surgery on the extent of glioblastoma resection. The question whether intraoperative ultrasound guided surgery enables complete tumor resection more often than standard surgery will be answered. In **Chapter 7** we evaluate the clinical feasibility of a wearable mixed reality device for planning glioblastoma surgery, presenting the first proof of concept.

Part III of this thesis consists of studies providing a postoperative evaluation of glioblastoma resection. In **Chapter 8** we assess the association between the resection of contrast enhancing and non-contrast enhancing parts of the tumor and survival in light of MGMT promoter methylation in a cohort of patients with newly diagnosed IDH wildtype glioblastoma. We answer the question whether complete resection is associated with improved survival in patients with molecularly defined glioblastoma. In relation to this, in **Chapter 9** we perform an international, multicenter, observational study, including over one thousand patients with a newly diagnosed IDH-wildtype glioblastoma, in which we develop and externally validate a survival prediction model. In **Chapters 10.1 and 10.2** we systematically review and assess the value of supratotal resection on patient survival, we present a meta-analysis and an editorial letter on this concept.

Finally, in **Chapter 11**, I provide a summary and an overall conclusion of this thesis.

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 Oncol. 2019;21(Suppl 5):v1-v100.
2. Chang SM, Parney IF, Huang W, Anderson FA, Jr., Asher AL, Bernstein M, et al. Patterns of care for adults with newly diagnosed malignant glioma. Jama. 2005;293(5):557-64.
3. Alexander BM, Cloughesy TF. Adult Glioblastoma. J Clin Oncol. 2017;35(21):2402-9.
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 Neuropathol. 2016;131(6):803-20.
5. van den Bent MJ. Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. Acta Neuropathol. 2010;120(3):297-304.
6. van den Bent MJ, Smits M, Kros JM, Chang SM. Diffuse Infiltrating Oligodendroglioma and Astrocytoma. J Clin Oncol. 2017;35(21):2394-401.
7. Louis DN, Wesseling P, Aldape K, Brat DJ, Capper D, Cree IA, et al. cIMPACT-NOW Update 6: New Entity and Diagnostic Principle Recommendations of the cIMPACT-Utrecht Meeting on Future CNS Tumor Classification and Grading. Brain Pathol. 2020 Jul;30(4):844-856.
8. Brat DJ, Aldape K, Colman H, Figrarella-Branger D, Fuller GN, Giannini C, et al. cIMPACT-NOW Update 5: Recommended Grading Criteria and Terminologies for IDH-mutant Astrocytomas. Acta Neuropathol . 2020 Mar;139(3):603-608. doi: 10.1007/s00401-020-02127-9. Epub 2020 Jan 29.
9. Lu C, Ward PS, Kapoor GS, Rohle D, Turcan S, Abdel-Wahab O, et al. IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature. 2012;483(7390):474-8.
10. Turcan S, Rohle D, Goenka A, Walsh LA, Fang F, Yilmaz E, et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature. 2012;483(7390):479-83.
11. Wang P, Wu J, Ma S, Zhang L, Yao J, Hoadley KA, et al. Oncometabolite D-2-Hydroxyglutarate Inhibits ALKBH DNA Repair Enzymes and Sensitizes IDH Mutant Cells to Alkylating Agents. Cell Rep. 2015;13(11):2353-61.
12. Cancer Genome Atlas Research N, Brat DJ, Verhaak RG, Aldape KD, Yung WK, Salama SR, et al. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med. 2015;372(26):2481-98.
13. Jenkins RB, Blair H, Ballman KV, Giannini C, Arusell RM, Law M, et al. A t(1;19)(q10;p10) mediates the combined deletions of 1p and 19q and predicts a better prognosis of patients with oligodendroglioma. Cancer Res. 2006;66(20):9852-61.
14. Kouwenhoven MC, Kros JM, French PJ, Biemond-ter Stege EM, Graveland WJ, Taphoorn MJ, et al. 1p/19q loss within oligodendroglioma is predictive for response to first line temozolomide but not to salvage treatment. Eur J Cancer. 2006;42(15):2499-503.

15. Cairncross JG, Ueki K, Zlatescu MC, Lisle DK, Finkelstein DM, Hammond RR, et al. Specific genetic predictors of chemotherapeutic response and survival in patients with anaplastic oligodendrogliomas. J Natl Cancer Inst. 1998;90(19):1473-9.
16. Esteller M, Garcia-Foncillas J, Andion E, Goodman SN, Hidalgo OF, Vanaclocha V, et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. N Engl J Med. 2000;343(19):1350-4.
17. Wick W, Platten M, Meisner C, Felsberg J, Tabatabai G, Simon M, et al. Temozolomide chemotherapy alone versus radiotherapy alone for malignant astrocytoma in the elderly: the NOA-08 randomised, phase 3 trial. Lancet Oncol. 2012;13(7):707-15.
18. Malmström A, Grønberg BH, Marosi C, Stupp R, Frappaz D, Schultz H, et al. Temozolomide versus standard 6-week radiotherapy versus hypofractionated radiotherapy in patients older than 60 years with glioblastoma: the Nordic randomised, phase 3 trial. Lancet Oncol. 2012;13(9):916-26.
19. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med. 2005;352(10):997-1003.
20. Eckel-Passow JE, Lachance DH, Molinaro AM, Walsh KM, Decker PA, Sicotte H, et al. Glioma Groups Based on 1p/19q, IDH, and TERT Promoter Mutations in Tumors. N Engl J Med. 2015;372(26):2499-508.
21. van den Bent MJ, Carpentier AF, Brandes AA, Sanson M, Taphoorn MJ, Bernsen HJ, et al. Adjuvant procarbazine, lomustine, and vincristine improves progression-free survival but not overall survival in newly diagnosed anaplastic oligodendrogliomas and oligoastrocytomas: a randomized European Organisation for Research and Treatment of Cancer phase III trial. J Clin Oncol. 2006;24(18):2715-22.
22. van den Bent MJ, Brandes AA, Taphoorn MJ, Kros JM, Kouwenhoven MC, Delattre JY, et al. Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: long-term follow-up of EORTC brain tumor group study 26951. J Clin Oncol. 2013;31(3):344-50.
23. Cairncross G, Wang M, Shaw E, Jenkins R, Brachman D, Buckner J, et al. Phase III trial of chemoradiotherapy for anaplastic oligodendroglioma: long-term results of RTOG 9402. J Clin Oncol. 2013;31(3):337-43.
24. Gevaert O, Mitchell LA, Achrol AS, Xu J, Echegaray S, Steinberg GK, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. Radiology. 2014;273(1):168-74.
25. Zinn PO, Colen RR. Imaging genomic mapping in glioblastoma. Neurosurgery. 2013;60 Suppl 1:126-30.
26. Kickingereder P, Bonekamp D, Nowosielski M, Kratz A, Sill M, Burth S, et al. Radiogenomics of Glioblastoma: Machine Learning-based Classification of Molecular Characteristics by Using Multiparametric and Multiregional MR Imaging Features. Radiology. 2016;281(3):907-18.
27. Pope WB. Genomics of brain tumor imaging. Neuroimaging Clin N Am. 2015;25(1):105-19.

28. Naeini KM, Pope WB, Cloughesy TF, Harris RJ, Lai A, Eskin A, et al. Identifying the mesenchymal molecular subtype of glioblastoma using quantitative volumetric analysis of anatomic magnetic resonance images. Neuro Oncol. 2013;15(5):626-34.
29. Smits M, van den Bent MJ. Imaging Correlates of Adult Glioma Genotypes. Radiology. 2017;284(2):316-31.
30. Gutman DA, Cooper LAD, Hwang SN, Holder CA. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set. Radiology. 2013.
31. Zhou H, Chang K, Bai HX, Xiao B, Su C, Bi WL, et al. Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas. J Neurooncol. 2019;142(2):299-307.
32. Eidel O, Burth S, Neumann JO, Kieslich PJ, Sahm F, Jungk C, et al. Tumor Infiltration in Enhancing and Non-Enhancing Parts of Glioblastoma: A Correlation with Histopathology. PLoS One. 2017;12(1):e0169292.
33. Verburg N, Koopman T, Yaqub MM, Hoekstra OS, Lammertsma AA, Barkhof F, et al. Improved detection of diffuse glioma infiltration with imaging combinations: a diagnostic accuracy study. Neuro Oncol. 2020;22(3):412-22.
34. Brown TJ, Brennan MC, Li M, Church EW, Brandmeir NJ, Rakszawski KL, et al. Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis. JAMA Oncol. 2016;2(11):1460-9.
35. Molinaro AM, Hervey-Jumper S, Morshed RA, Young J, Han SJ, Chunduru P, et al. Association of Maximal Extent of Resection of Contrast-Enhanced and Non-Contrast-Enhanced Tumor With Survival Within Molecular Subgroups of Patients With Newly Diagnosed Glioblastoma. JAMA Oncol. 2020:10.1001/jamaoncol.2019.6143.
36. Li YM, Suki D, Hess K, Sawaya R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? J Neurosurg. 2016;124(4):977-88.
37. Grabowski MM, Recinos PF, Nowacki AS, Schroeder JL, Angelov L, Barnett GH, et al. Residual tumor volume versus extent of resection: Predictors of survival after surgery for glioblastoma. J Neurosurg. 2014;121(5):1115-23.
38. 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. J Neurosurg. 2001;95(2):190-8.
39. Noorbakhsh A, Tang JA, Marcus LP, McCutcheon B, Gonda DD, Schallhorn CS, et al. Gross-total resection outcomes in an elderly population with glioblastoma: A SEER-based analysis. Clinical article. J Neurosurg. 2014;120(1):31-9.
40. Pessina F, Navarria P, Cozzi L, Ascolese AM, Simonelli M, Santoro A, et al. Maximize surgical resection beyond contrast-enhancing boundaries in newly diagnosed glioblastoma multiforme: is it useful and safe? A single institution retrospective experience. J Neuro-Oncol. 2017;135(1):129-39.

41. Eyüpoglu IY, Hore N, Merkel A, Buslei R, Buchfelder M, Savaskan N. Supra-complete surgery via dual intraoperative visualization approach (DiVA) prolongs patient survival in glioblastoma. Oncotarget. 2016;7(18):25755-68.
42. de Leeuw CN, Vogelbaum MA. Supratotal resection in glioma: a systematic review. Neuro Oncol. 2019;21(2):179-88.
43. Jenkinson MD, Barone DG, Bryant A, Vale L, Bulbeck H, Lawrie TA, et al. Intraoperative imaging technology to maximise extent of resection for glioma Review. Cochrane Database Syst Rev. 2018;1:CD012788.
44. Stummer W, Pichlmeier U, Meinel T, Wiestler OD. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. The lancet oncology. 2006.
45. Senft C, Bink A, Franz K, Vatter H, Gasser T, Seifert V. Intraoperative MRI guidance and extent of resection in glioma surgery: a randomised, controlled trial. Lancet Oncol. 2011;12(11):997-1003.
46. Unsgaard G, Ommedal S, Muller T, Gronningsaeter A, Nagelhus Hernes TA. Neuronavigation by intraoperative three-dimensional ultrasound: initial experience during brain tumor resection. Neurosurgery. 2002;50(4):804-12; discussion 12.
47. Hollon TC, Pandian B, Adapa AR, Urias E, Save AV, Khalsa SSS, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. Nat Med. 2020;26(1):52-8.
48. Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJB, Janzer RC, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol. 2009;10(5):459-66.
49. Stupp R, Mason WP, Van Den Bent MJ, Weller M, Fisher B, Taphoorn MJB, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. New Engl J Med. 2005;352(10): 987-96.
50. Marenco-Hillembrand L, Wijesekera O, Suarez-Meade P, Mampre D, Jackson C, Peterson J, et al. Trends in glioblastoma: outcomes over time and type of intervention: a systematic evidence based analysis. J Neurooncol. 2020;147(2):297-307.
51. Drumm MR, Dixit KS, Grimm S, Kumthekar P, Lukas RV, Raizer JJ, et al. Extensive brainstem infiltration, not mass effect, is a common feature of end-stage cerebral glioblastomas. Neuro Oncol. 2020;22(4):470-9.
52. Silbergeld DL, Rostomily RC, Alvord EC Jr. The cause of death in patients with glioblastoma is multifactorial: clinical factors and autopsy findings in 117 cases of supratentorial glioblastoma in adults. J Neurooncol. 1991;10(2):179–185.

PART I

Preoperative imaging

Chapter 2

Topographical Mapping of 436 Newly Diagnosed IDH Wildtype Glioblastoma with vs. without MGMT Promoter Methylation

Fatih Incekara*, Sebastian van der Voort*, Hendrikus J. Dubbink, Peggy N. Atmodimedjo, Rishi Nandoe Tewarie, Geert Lycklama, Arnaud J.P.E. Vincent, Johan M. Kros, Stefan Klein, Martin van den Bent, Marion Smits

Front. Oncol. 2020 May 12: 10:596.

* contributed equally

Abstract

Background

O^6-methylguanine-methyltransferase (MGMT) promoter methylation and isocitrate dehydrogenase (IDH) mutation status are important prognostic factors for patients with glioblastoma. There are conflicting reports about a differential topographical distribution of glioblastoma with vs. without MGMT promoter methylation, possibly caused by molecular heterogeneity in glioblastoma populations. We initiated this study to re-evaluate the topographical distribution of glioblastoma with vs. without MGMT promoter methylation in light of the updated WHO 2016 classification.

Methods

Pre-operative T2-weighted/FLAIR and post-contrast T1-weighted MRI scans of patients aged 18 year or older with IDH wildtype glioblastoma were collected. Tumors were semi-automatically segmented and the topographical distribution between glioblastoma with vs. without MGMT promoter methylation was visualized using frequency heatmaps. Then voxel-wise differences were analyzed using permutation testing with Threshold Free Cluster Enhancement.

Results

Four hundred thirty-six IDH wildtype glioblastoma patients were included; 211 with and 225 without MGMT promoter methylation. Visual examination suggested that when compared with MGMT unmethylated glioblastoma, MGMT methylated glioblastoma were more frequently located near bifrontal and left occipital periventricular area and less frequently near the right occipital periventricular area. Statistical analyses, however, showed no significant difference in topographical distribution between MGMT methylated vs. MGMT unmethylated glioblastoma.

Conclusion

This study re-evaluated the topographical distribution of MGMT promoter methylation in 436 newly diagnosed IDH wildtype glioblastoma, which is the largest homogenous IDH wildtype glioblastoma population to date. There was no statistically significant difference in anatomical localization between MGMT methylated vs. unmethylated IDH wildtype glioblastoma.

Introduction

Patients with glioblastoma have a poor prognosis with a median overall survival of 15 months, despite standard of care consisting of safe and maximal surgical resection followed by chemo- and/or radiotherapy.[1] This prognosis varies based on factors such as age, Karnofsky Performance Status, extent of resection and molecular markers, in particular isocitrate dehydrogenase (IDH) mutation and O^6-methylguanine-methyltransferase (MGMT) promoter methylation status.[2]

MGMT is a DNA repair enzyme, which is expressed by the MGMT gene located on chromosome 10q26. Promoter methylation of this gene reduces MGMT protein expression and consequently decreases DNA repair and increases alkylating chemotherapy induced tumor death. Therefore, patients with MGMT methylated glioblastoma are more sensitive to neo-adjuvant temozolomide than those without MGMT methylated glioblastoma. MGMT is methylated in approximately 50% of patients with newly diagnosed glioblastoma.[3]

There are conflicting results in the published literature on a possible differential topographical distribution of glioblastoma with vs. without MGMT promoter methylation.[4] Ellingson et al. suggested that when compared with those without MGMT promoter methylation, glioblastoma with methylation are more frequently located in the left temporal lobe and less frequently in the right temporal lobe.[5] However, other studies found the reverse lateralization pattern[6] or did not find any lateralization at all.[7-9] These conflicting results could be ascribed to heterogeneity of molecular subtypes of glioblastoma in the studied populations, for instance when IDH wildtype glioblastoma are mixed with the genetically and prognostically distinct IDH mutated glioblastoma, or to variation in statistical methods that were used across studies. Therefore, the question whether glioblastoma with vs. without MGMT promoter methylation have a different anatomical localization remains unanswered. In light of the updated WHO 2016 classification[10], a molecularly homogenous glioblastoma population must be used to re-evaluate the topographical distribution of MGMT methylated vs. unmethylated glioblastoma.

Therefore, we have initiated this study to re-evaluate the topographical distribution of glioblastoma with and vs. without MGMT promoter methylation in the largest homogenous IDH wildtype glioblastoma population to date.

METHODS

Patient inclusion

All consecutive patients aged 18 years or older newly diagnosed with a contrast-enhancing and histopathologically confirmed glioblastoma IDH wildtype who underwent tumor resection or biopsy between January 2011 and May 2018 at the Erasmus MC, University Medical Center Rotterdam or Haaglanden MC were retrospectively included in

this study. Patients were eligible if pre-operative T2-weighted/fluid-attenuated inversion recovery (FLAIR) and post-contrast T1-weighted MRI scans as well as molecular data on IDH mutation and MGMT methylation status were available. Recurrent glioblastoma or confirmed IDH mutated glioblastoma were excluded. The study design was approved by the Medical Ethical Committee of Erasmus MC and Haaglanden MC. The study was performed in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Image acquisition, tumor segmentation and registration

From clinical pre-operative MRI scans, which were obtained according to clinical brain tumor protocols on either a 1.5T or 3.0T scanner, T2-weighted/FLAIR and post-contrast T1-weighted images were collected. For glioblastoma segmentation, we first imported both the post-contrast T1-weighted and T2-weighted/FLAIR scans into BrainLab (BrainLab, Feldkirchen, Germany, version 2.1.0.15). We semi-automatically segmented all tumor-related contrast-enhancement (including the central necrotic part, if present) using the SmartBrush tool in Brainlab Elements and manually adapted the segmentation if needed. We then used the T2-weighted/FLAIR scan to semi-automatically segment all tumor-related non-enhancing hyperintense abnormalities (extra-lesional hemorrhage were excluded).

All tumor segmentations were then registered to the Montreal Neurological Institute (MNI) International Consortium for Brain Mapping 152 nonlinear atlas. The post-contrast T1-weighted scans were registered to the T1-weighted atlas and the T2-weighted/FLAIR scans to the T2-weighted atlas. Registration was done using SimpleElastix (version 72b7e81), based on a mutual information metric using an affine registration.[11] The resulting transformation parameters were used to transform the 3D segmentations to the atlas space. Registration results were visually checked to ensure that for all cases the registered masks lay entirely within the brain mask of the atlas. No adjustments were made to the initial registration settings for individual patients. We created voxel-wise frequency maps for all glioblastoma combined, and frequency difference maps of glioblastoma with versus without MGMT promoter methylation.

Molecular analysis

Tumor tissue samples were obtained from patients through surgical resection or biopsy. Histopathological examination was performed by neuropathologists. DNA was extracted from microdissected FFPE tissue fragments by proteinase K digestion for 16 h at 56 C in the presence of 5 % Chelex 100 resin and used after inactivation of proteinase K and removal of cell debris and the Chelex resin. IDH mutational analysis was assessed with Sanger sequencing of PCR-amplified fragments from IDH1 and IDH2 mutational hotspots, essentially as previously described.[12] M13-tailed primers for PCR amplification of IDH1 codon 132 were forward 5'-TGTAAA ACGACGGCCAGTCTCCTGATGAGAAGAGGGTTG-3' and reverse 5'–CAGGAAACAGCTATGACCCATT CTCTGGTTTTCGCATGCAAAATCACATTATTGCC-3'. After

initial denaturation at 95 C for 3 min, 35 cycles of 95 C for 30 s, 60 C for 45 s, and 72 C for 45 s were performed, followed by 10 min at 72 C. Subsequent sequence analyses of the PCR products was carried out with M13 forward and reverse primers on an 3730 XL Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Targeted NGS was performed by semiconductor sequencing with the Ion Torrent platform using supplier's materials and protocols (Thermo Fisher Scientific) with a dedicated panel for detection of glioma-specific aberrations, including IDH1 and IDH2 hot spot mutations essentially as previously described.[13] Library and template preparations were performed consecutively with the AmpliSeq Library Kit 2.0-384 LV and the Ion 510/520/530 Chef kit. Sequencing was performed on a 530 or 540 chip with the Ion S5 XL system. Data was analyzed with the Torrent variant caller (Thermo Fisher Scientific) and variants were annotated in a local Galaxy pipeline using ANNOVAR. Details of the glioma panel are described in the supplementary data of Dubbink et al.[13]

MGMT promoter methylation status was assessed by methylation-specific PCR essentially as described by Esteller et al.[14] Bisulfite conversion and subsequent purification is performed with the EZ DNA Methylation-Gold Kit (Zymo Research) according to the supplier's protocol. Methylation-specific PCR was performed with primers specific for either methylated or the modified unmethylated DNA. Converted primer sequences for unmethylated DNA were forward 5'-TTTGTGTTTTGATGTTTGTAGGTTTTTGT-3' and reverse 5'-AACTCCACACTCTTCCAAAAACAAAACA-3', and for the methylated reaction, forward 5'-TTTCGACGTTCGTAGGTTTTCGC-3' and reverse 5'-GCACTCTTCCGAAAACGAAACG-3 . PCR was performed after initial denaturation at 95 C for 5 min by 35 cycles of 92 C for 45 s, 59 C for 65 s, and 72 C for 45 s, followed by 7 min at 72 C. Five microliters of each 15 µl methylation-specific PCR product was loaded onto a 1.5 % agarose gel stained with GelRed (Biotium) and examined under ultraviolet illumination. SW48 cell line DNA and tonsil DNA was used as a positive control for methylated and unmethylated alleles of MGMT, respectively. Controls without DNA were used for each set of methylation-specific PCR assays.

Statistical analysis

We first tested the differences between pre-operative enhancing and non-enhancing tumor volumes as well as their ratio with the Kruskal-Wallis test. We mapped the anatomical localization of all MGMT methylated and unmethylated glioblastoma by iterating over all voxels in the MNI atlas and counting the number of tumor frequencies for each group in each voxel. To test for differences in spatial distribution between glioblastoma with vs. without MGMT promoter methylation, we assessed the cluster-wise significance at the voxel-level between distributions, using permutation testing with Threshold Free Cluster Enhancement [15] in the software package "FSL Randomize" (version 5.0.9, using 10,000 permutations).[16] This approach corrects p-values for the family-wise error in testing multiple voxels, considering a corrected p-value of <0.05 as statistically significant.

Results

In total, 769 patients with newly diagnosed, contrast enhancing glioblastoma were screened, of which we excluded 333 patients: 22 were excluded due to IDH mutation and 311 were excluded due to insufficient or missing molecular data on IDH mutation or MGMT methylation status. Final analysis included 436 patients with IDH wildtype glioblastoma (see flowchart, Supplementary Material); 211 with and 225 without MGMT promoter methylation. 340 patients had undergone a surgical tumor resection and 96 a diagnostic biopsy. In all patients pre-operative post-contrast T1-weighted MRI scans were available; in 90 patients T2-weighted FLAIR scans and in 346 patients T2-weighted scans were available. When compared with MGMT unmethylated glioblastoma, MGMT methylated glioblastoma had a significantly higher ratio of non-enhancing versus contrast-enhancing volume (2.09 (inter quartile range 2.6) and 2.5 (inter quartile range 3.3), p=0.045, respectively). Patient and tumor characteristics are further presented in Table 1.

Table 1. Patient and tumor characteristics.

Characteristics	**n**	**%**	
all patients	436	100	
Sex			
Male	276	63.3	
Female	160	36.7	
Age	x	x	
≤ 65	227	52.1	
> 65	209	47.9	
Mean, years (SD)	61.5 (16.2)		
Karnofsky Performance Status			
≤ 70	142	32.6	
> 70	294	67.4	
Mean (SD)	80 (12.5)		
Pre-operative MRI scans			
T1 post-contrast	436	100	
T2-weighted	346	79.4	
T2-weighted FLAIR	90	20.6	
Neurosurgical procedure			
Resection	340	78.0	
Biopsy	96	22.0	
	MGMT promotor		
Pre-operative volume, median cm³ (IQR)	**Methylated** 211 (48.4%)	**Unmethylated** 225(51.6%)	p- value
Contrast-enhancing	30.1 (39.5)	35 (45.8)	.130
Non-enhancing	75.5 (105.0)	65.5 (84.2)	.338
Non-enhancing/contrast-enhancing Ratio	2.5 (3.3)	2.09 (2.6)	.045

SD standard deviation, IQR Inter Quartile Range, CE contrast enhancement, FLAIR Fluid-attenuated inversion recovery.

Topographical mapping of 436 IDH wildtype glioblastoma

For visual inspection, heatmaps based on post-contrast T1-weighted and T2-weighted/FLAIR segmentations were created for all 436 patients combined (Figures 1 and 2), as well as frequency difference maps between MGMT methylated vs. unmethylated glioblastoma (Figure 3). Visual inspection of maps in Figure 1 suggests that glioblastoma were most frequently located in the right temporal, insular and parietal area, and near the periventricular area both frontally and occipitally. Visual inspection of Figures 2 and 3 indicates that when compared with MGMT unmethylated glioblastoma, methylated glioblastoma were more frequently located near bifrontal and left occipital periventricular area (up to 6.5% frequency difference) and less frequently near the right occipital periventricular area (up to 9.1% frequency difference).

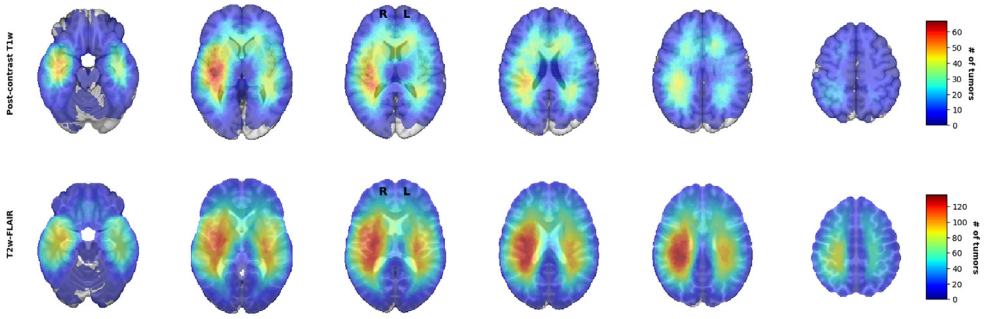


Figure 1. Heatmaps of all 436 IDH wildtype GBM.

To test whether this difference was statistically significant, voxel-wise analyses of both the post-contrast T1-weighted and T2-weighted/FLAIR segmentation heatmaps were performed. Although statistical analysis of the post-contrast T1-weighted scans marked a region near the right occipital periventricular area as a potentially discriminating area between MGMT methylated vs. unmethylated glioblastoma, this difference was not statistically significant (Figure 4, together with corresponding p-values). This figure in fact shows that not any statistically significantly discriminating brain area between MGMT methylated and unmethylated glioblastoma could be found. Scroll-through video clips for visual inspection of all topographic maps are publicly available as Supplementary Material.

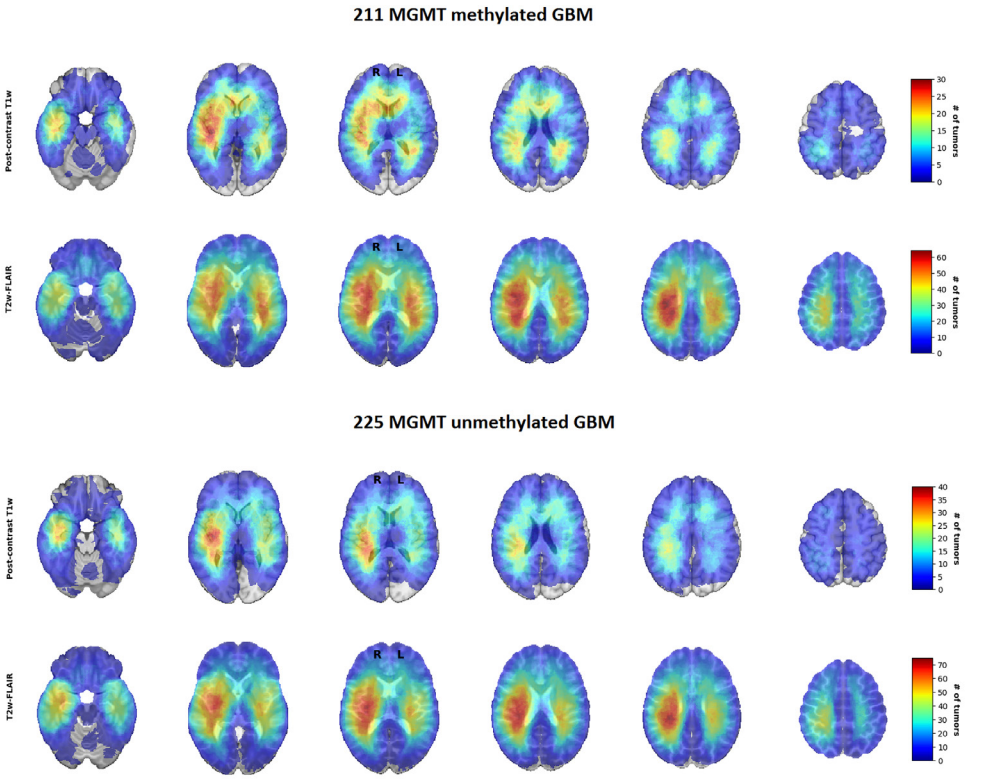


Figure 2. Heatmaps of MGMT methylated (N=211) and unmethylated (N=225) GBM.

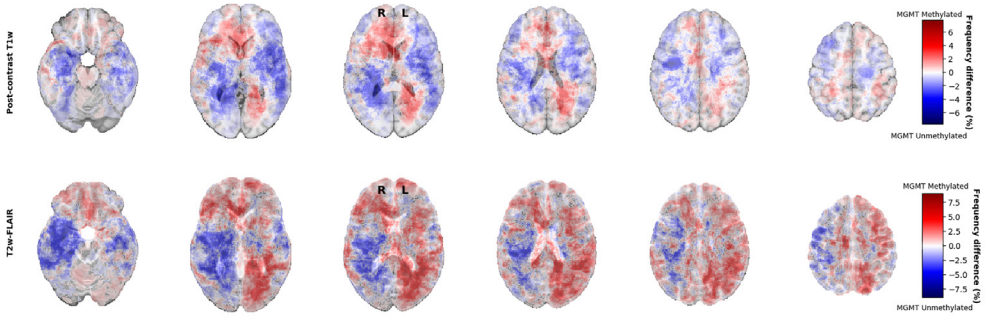


Figure 3. Frequency difference maps between MGMT methylated (N=211) and unmethylated (N=225) GBM.

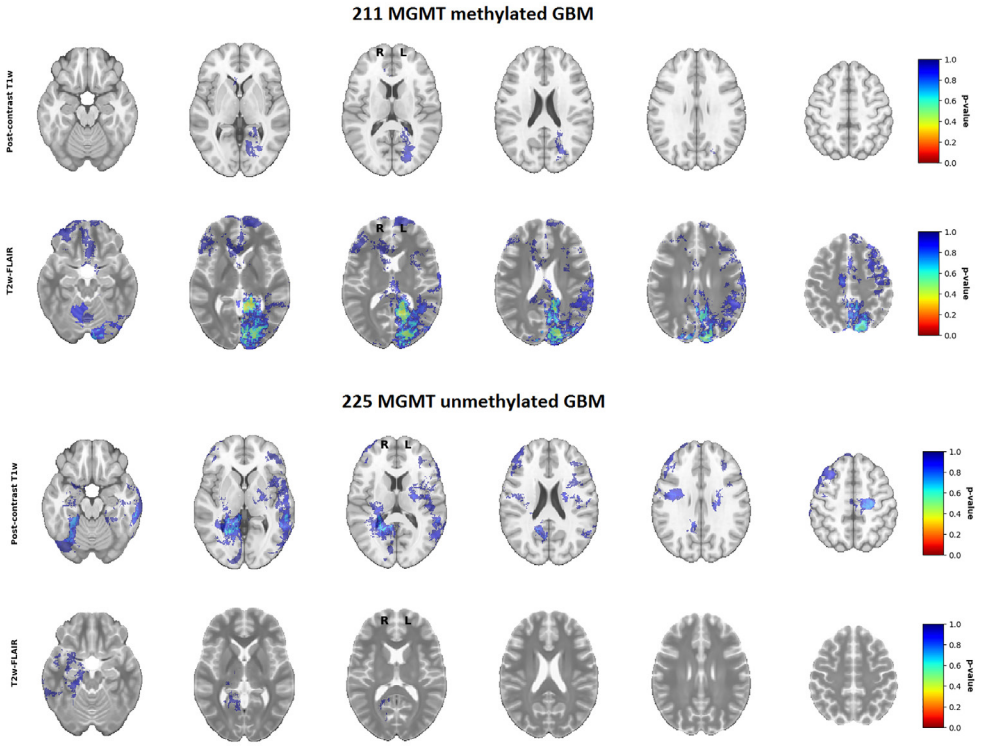


Figure 4. P-value maps of MGMT methylated (N=211) and unmethylated (N=225) GBM.

Discussion

This study voxel-wise analyzed post-contrast T1-weighted and T2-weighted/FLAIR heatmaps and showed that there was no statistically significant difference in anatomical localization between MGMT methylated vs. unmethylated IDH wildtype glioblastoma.

The primary reason to initiate this study was to re-evaluate the anatomic localization of MGMT methylated vs. unmethylated glioblastoma in light of the updated WHO 2016 classification era following conflicting reports on this topic.[4] Ellingson et al. (2013) reported that glioblastoma with MGMT methylation were lateralized to the left hemisphere (temporal lobe) and that those without were lateralized to the right hemisphere(17), which was in line with their previous article (2012) and in which they included a substantial portion of their previously studied glioblastoma population.[5] However, in contrast to these findings there are also studies that found the reverse pattern of hemispheric lateralization, in which glioblastoma with MGMT methylation were located more frequently in the right hemisphere, while those without MGMT methylation lateralized to the left hemisphere.[6] Additionally, there are conflicting reports on lobar distribution, in which glioblastoma with MGMT methylation were more frequently located in the parietal

and occipital lobes, while those without were located more frequently in the temporal lobes.[8] A recent study suggested after qualitative analyses that subventricular zones were more frequently spared with MGMT methylated glioblastoma, but found no difference in hemispheric lateralization between glioblastoma with and without MGMT promoter methylation.[9] Finally, there are also studies that report no differences in localization between glioblastoma with and without MGMT methylation,[7, 18] in concordance with the findings of our study.

These conflicting results in the literature can potentially be ascribed to two methodological issues. First, inconsistencies may arise from variations in glioblastoma patient populations across studies, many of which were performed in the pre-WHO 2016 classification era when the impact of molecular subtyping of glioblastoma according to IDH mutation status was less of a consideration.[10] Ellingson et al. (2013) included a series of 507 de novo glioblastoma with mixed IDH subtypes, including 366 IDH wildtype, 34 IDH mutated glioblastoma and also 107 glioblastoma without data on IDH mutation status.(17) Moreover, the majority of the studies did not report the IDH mutation status of included glioblastoma.[5, 6, 8, 18]

Mixing molecular subtypes or not knowing IDH mutation status of glioblastoma is undesirable when assessing topographical distribution of molecular subtypes,[10] since it is now known that IDH mutated glioblastoma represent a distinct molecular subtype of glioblastoma from a distinct precursor lesion which have a predominantly frontal lobe involvement when compared with IDH wildtype glioblastoma.[19] This topographic link between IDH mutation and MGMT methylation was also suggested by Ellingson et al. (2013) by demonstrating that IDH mutated and MGMT methylated glioblastoma were indeed more frequently localized in the frontal lobe.[17] This has not only been demonstrated in glioblastoma, but also in non-contrast enhancing low grade glioma in which IDH mutated low grade glioma (both oligodendroglioma and astrocytoma) were more frequently located in the frontal lobes, while non-contrast enhancing IDH wildtype astrocytoma were more frequently located in the basal ganglia of the right hemisphere.[20] This topographical link thus suggests IDH mutation status as (confounding) factor between MGMT methylation status and localization. Therefore, studies must be conducted based on homogeneous tumor populations with respect to IDH mutational status. This hypothesis was recently supported by Roux et al, who assessed a homogenous IDH wildtype glioblastoma population (n = 392) and found no difference in localization between glioblastoma with and without MGMT methylation, in line with our study.[21]

Second, the conflicting results in the literature may arise from different statistical methods that were used across studies. Studies often investigated the anatomic localization of glioblastoma with and without MGMT promoter methylation with visual examination, qualitatively, without a statistical, voxel-wise quantitative analysis.[7-9,18] Ellingson et al. (2013) used frequency difference maps to demonstrate that MGMT methylated glioblastoma were more frequently localized in the left temporal lobe.[17]

Using similar frequency difference maps, we also found topographical differences, which indicated that when compared with MGMT unmethylated glioblastoma, MGMT methylated glioblastoma were more frequently localized near bifrontal and right occipital periventricular area and less frequently near the right occipital periventricular area. However, we showed that these apparent differences did not survive rigorous statistical testing. Ellingson et al. report the use of 'Analysis of Differential Involvement' for their statistical analysis, which is based on the Fisher exact test.[5] We used 'FSL randomize', which is different from the Fisher exact test because it does not make any assumptions about the underlying distribution of the variables.[16] Another methodological difference can be found in the correction for multiple comparisons. Ellingson et al. used random permutations based on Bullmore et al. instead of the more recently proposed and widely accepted method of doing random permutations employed in 'FSL randomize' based on Smith et al.[15,22] Furthermore, the method by Bullmore et al. requires a user-defined threshold for clustering, which can impact the results substantially.[22] Instead, we used 'Threshold Free Cluster Enhancement', which does not require thresholding to determine the clusters, and which has been shown to have a higher sensitivity compared to other methods.[15] Our stringent methodology of rigorous statistical testing and applying new insights in glioblastoma molecular subtyping to a large studied patient population are the strengths of our study.

Limitations

The main limitation of this study is its retrospective design, which may have introduced selection and confounding biases. Selection bias may occur when patients who receive diagnostic biopsies are excluded from analysis, since these tumors are often large, multifocal, located deep within the basal ganglia, or crossing midline. This may skew the results on tumor localization of glioblastoma, which is our main outcome. We have therefore attempted to limit this bias first by consecutive inclusion of all glioblastoma patients operated upon between 2011 and 2018 in our cohort, including diagnostic biopsies. In addition, it is known that tumor localization is associated with IDH mutation status, with IDH mutated tumors located more frequently in the frontal lobes, as mentioned earlier.[19] Since IDH mutation status is both associated with tumor localization and MGMT methylation status, it may function as a confounding factor. We therefore have also attempted to limit this potential bias by excluding all IDH mutated tumors. Another limitation is that we included patients from two medical centers from a period of over seven years. This introduced variation of MRI scan protocols such as magnet strength, voxel size and slice thickness, which consequently may have negatively influenced registration accuracy and anatomical localization. Such registration inaccuracies can however be considered minor relative to the size of the tumor and it is therefore unlikely that our results were significantly impacted by scanner variations. Additionally, tumor volume assessment on these MRI scans were performed by one observer without

confirmation of a second, independent assessor. This may have introduced some degree of information bias. We have attempted to limit this bias during volumetric assessment by blinding the assessor for patients' clinical and molecular characteristics. It is known that both the inter and intra-observer agreement for pre-operative tumor volumes in glioblastoma is relatively high.[23] Finally, it should be noted that the known intertest variability is a limitation of MGMT analyses, as assays used in other studies may produce slightly different MGMT methylation results.[24] This may partially explain the variety in the proportion of MGMT methylated tumors reported in literature.

Conclusion

In the largest homogenous IDH wildtype glioblastoma population to date, we showed that visual appearance of differences could not be confirmed with rigorous voxel-wise statistical testing and thus that there is no statistical difference in anatomical localization between IDH wildtype glioblastoma with vs. without MGMT promoter methylation.

References

1. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med* (2005) 352(10):987-96. PubMed PMID: 15758009.
2. Gessler F, Bernstock JD, Braczynski A, Lescher S, Baumgarten P, Harter PN, et al. Surgery for Glioblastoma in Light of Molecular Markers: Impact of Resection and MGMT Promoter Methylation in Newly Diagnosed IDH-1 Wild-Type Glioblastomas. *Neurosurgery* (2019) 84(1):190-7. PubMed PMID: 29617848.
3. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* (2005) 352(10):997-1003. PubMed PMID: 15758010.
4. Smits M, van den Bent MJ. Imaging Correlates of Adult Glioma Genotypes. *Radiology* (2017) 284(2):316-31. PubMed PMID: 28723281.
5. Ellingson BM, Cloughesy TF, Pope WB, Zaw TM, Phillips H, Lalezari S, et al. Anatomic localization of O6-methylguanine DNA methyltransferase (MGMT) promoter methylated and unmethylated tumors: a radiographic study in 358 de novo human glioblastomas. *Neuroimage* (2012) 59(2):908-16. PubMed PMID: 22001163.
6. Wang Y, Fan X, Zhang C, Zhang T, Peng X, Li S, et al. Anatomical specificity of O6-methylguanine DNA methyltransferase protein expression in glioblastomas. *J Neurooncol* (2014) 120(2):331-7. PubMed PMID: 25031184.
7. Carrillo JA, Lai A, Nghiemphu PL, Kim HJ, Phillips HS, Kharbanda S, et al. Relationship between tumor enhancement, edema, IDH1 mutational status, MGMT promoter methylation, and survival in glioblastoma. *AJNR Am J Neuroradiol* (2012) 33(7):1349-55. PubMed PMID: 22322613.
8. Eoli M, Menghi F, Bruzzone MG, De Simone T, Valletta L, Pollo B, et al. Methylation of O6-methylguanine DNA methyltransferase and loss of heterozygosity on 19q and/or 17p are overlapping features of secondary glioblastomas with prolonged survival. *Clin Cancer Res* (2007) 13(9):2606-13. PubMed PMID: 17473190.
9. Han Y, Yan LF, Wang XB, Sun YZ, Zhang X, Liu ZC, et al. Structural and advanced imaging in predicting MGMT promoter methylation of primary glioblastoma: a region of interest based analysis. *BMC Cancer* (2018) 18(1):215. PubMed PMID: 29467012.
10. 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 Neuropathol* (2016) 131(6):803-20. PubMed PMID: 27157931.
11. Marstal K, Berendsen F, Staring M, Klein S, editors. SimpleElastix: A User-Friendly, Multi-lingual Library for Medical Image Registration. *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; 2016 26 June-1 July 2016.
12. van den Bent MJ, Hartmann C, Preusser M, Ströbel T, Dubbink HJ, Kros JM, et al. Interlaboratory comparison of IDH mutation detection. *J Neurooncol* (2013) 112(2):173-8. Epub 2013/01/29. doi: 10.1007/s11060-013-1056-z. PubMed PMID: 23358936.

13. Dubbink HJ, Atmodimedjo PN, Kros JM, French PJ, Sanson M, Idbaih A, et al. Molecular classification of anaplastic oligodendroglioma using next-generation sequencing: a report of the prospective randomized EORTC Brain Tumor Group 26951 phase III trial. *Neuro Oncol* (2016) 18(3):388-400. PubMed PMID: 26354927.
14. Esteller M, Garcia-Foncillas J, Andion E, Goodman SN, Hidalgo OF, Vanaclocha V, et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med* (2000) 343(19):1350-4. doi: 10.1056/nejm200011093431901. PubMed PMID: 11070098.
15. Smith SM, Nichols TE. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* (2009) 44(1):83-98. PubMed PMID: 18501637.
16. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. *Neuroimage* (2014) 92:381-97. PubMed PMID: 24530839.
17. Ellingson BM, Lai A, Harris RJ, Selfridge JM, Yong WH, Das K, et al. Probabilistic radiographic atlas of glioblastoma phenotypes. *Am J Neuroradiol* (2013) 34(3):533-40. doi: 10.3174/ajnr.A3253.
18. Drabycz S, Roldan G, de Robles P, Adler D, McIntyre JB, Magliocco AM, et al. An analysis of image texture, tumor location, and MGMT promoter methylation in glioblastoma using magnetic resonance imaging. *Neuroimage* (2010) 49(2):1398-405. PubMed PMID: 19796694.
19. Lai A, Kharbanda S, Pope WB, Tran A, Solis OE, Peale F, et al. Evidence for sequenced molecular evolution of IDH1 mutant glioblastoma from a distinct cell of origin. *J Clin Oncol* (2011) 29(34):4482-90. PubMed PMID: 22025148.
20. Chaichana KL, Halthore AN, Parker SL, Olivi A, Weingart JD, Brem H, et al. Factors involved in maintaining prolonged functional independence following supratentorial glioblastoma resection: Clinical article. *J Neurosurg* (2011) 114(3):604-12. doi: 10.3171/2010.4.jns091340.
21. Roux A, Roca P, Edjlali M, Sato K, Zanello M, Dezamis E, et al. MRI Atlas of IDH Wild-Type Supratentorial Glioblastoma: Probabilistic Maps of Phenotype, Management, and Outcomes. *Radiology* (2019):190491. PubMed PMID: 31592732.
22. Bullmore ET, Suckling J, Overmeyer S, Rabe-Hesketh S, Taylor E, Brammer MJ. Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Trans Med Imaging* (1999) 18(1):32-42. PubMed PMID: 10193695.
23. Kubben PL, Postma AA, Kessels AGH, Van Overbeeke JJ, Van Santbrink H. Intraobserver and interobserver agreement in volumetric assessment of glioblastoma multiforme resection. *Neurosurgery* (2010) 67(5):1329-34. doi: 10.1227/NEU.0b013e3181efbb08.
24. Wick W, Weller M, van den Bent M, Sanson M, Weiler M, von Deimling A, et al. MGMT testing--the challenges for biomarker-based glioma treatment. *Nat Rev Neurol* (2014) 10(7):372-85. Epub 2014/06/10. doi: 10.1038/nrneurol.2014.100. PubMed PMID: 24912512.

Supplementary materials

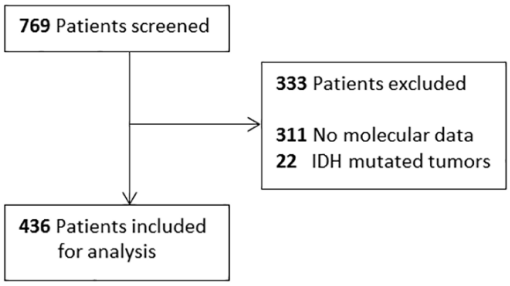


Figure S1. Flowchart.

The following heatmaps are online available:

- Heatmaps all GBM (T1w and T2w).gif
- Heatmaps methylated and unmethylated (T1w and T2w).gif
- Frequency difference maps (T1w and T2w).gif
- P value maps methylated and unmethylated (T1w and T2w).gif

Chapter 3

Predicting the 1p/19q co-deletion status of presumed low grade glioma with an externally validated machine learning algorithm

Sebastian van der Voort*, Fatih Incekara*, Maarten M.J. Wijnenga, Georgios Kapas, Mayke Gardeniers, Joost W. Schouten, Martijn P.A. Starmans, Rishi Nandoe Tewarie, Geert J. Lycklama, Pim J. French, Hendrikus J. Dubbink, Martin van den Bent, Arnaud J.P.E. Vincent, Wiro J. Niessen, Stefan Klein, Marion Smits

Clin Cancer Res. 2019 Dec 15;25(24):7455-7462.

* contributed equally

Abstract

Background

Patients with 1p/19q co-deleted low-grade glioma (LGG) have longer overall survival and better treatment response than patients with 1p/19q intact tumors. Therefore, it is relevant to know the 1p/19q status. To investigate whether the 1p/19q status can be assessed prior to tumor resection, we developed a machine learning algorithm to predict the 1p/19q status of presumed LGG based on preoperative magnetic resonance imaging (MRI).

Methods

Preoperative brain MRI scans from 284 patients who had undergone biopsy or resection of presumed LGG were used to train a support vector machine algorithm. The algorithm was trained based on features extracted from T1-weighted and T2-weighted MRI scans, and on patient age and sex. The performance of the algorithm compared to tissue diagnosis was assessed on an external validation dataset of MRI scans from 129 LGG patients from The Cancer Imaging Archive (TCIA). Four clinical experts also predicted the 1p/19q status of the TCIA MRI scans.

Results

The algorithm achieved an area under the curve (AUC) of 0.72 in the external validation dataset. The algorithm had a higher predictive performance than the average of the neurosurgeons (AUC 0.52), but lower than that of the neuroradiologists (AUC 0.81). There was a wide variability between clinical experts (AUC 0.45-0.83).

Conclusion

Our results suggest that our algorithm can non-invasively predict the 1p/19q status of presumed LGG with a performance that on average outperformed the oncological neurosurgeons. Evaluation on an independent dataset indicates that our algorithm is robust and generalizable.

Introduction

Low grade glioma (LGG) are primary brain tumors that originate from glial cells. The World Health Organization (WHO) 2016 criteria recognize three subtypes based on molecular and histological features:[1] diffuse IDH wildtype astrocytoma (IDH wildtype, 1p/19q intact),[2] diffuse IDH mutant astrocytoma (IDH mutated, 1p/19q intact); and (3) oligodendroglioma (IDH mutated, 1p/19q co-deleted).[1,2]

Studies have shown that the distinction between these three categories is clinically relevant in terms of prognosis and management: in patients treated with optimal surgical resection followed by radiation therapy with or without chemotherapy, median survival is longest of those with oligodendroglioma.[3,4] Additionally, studies have suggested that residual tumor has a more negative impact on survival in 1p/19q intact, IDH mutated astrocytoma than on 1p/19q co-deleted, IDH mutated oligodendrogliomas.[5,6] Therefore, the ability to predict the molecular subtypes of LGG at an early stage could provide better guidance of risk-benefit assessment and clinical decision-making.

The recent shift from histopathology-based glioma classification to the molecular subtype-based WHO 2016 classification gave rise to neuro-oncological radiogenomics research in which features seen on preoperative magnetic resonance imaging (MRI) scans are used to predict the genetic mutation status of glioma.[7-9] Features such as frontal tumor localization, indistinct tumor borders, heterogeneous signal intensity on T2-weighted images, and both cortical and subcortical tumor infiltration all suggest the presence of 1p/19q co-deletion.[7]

One way of linking MRI features to 1p/19q co-deletion is through machine learning. While several studies have applied this method to datasets of patients with high grade glioma, few studies have developed radiogenomics methodology in LGG.[10-15] Of the ones that have, most have not used an independent test set and, therefore, it is difficult to estimate their actual performance in the real-world clinical setting.[10,11,13,14] Lu et al.[12] did use an independent test set, but this set contained a very limited number of LGG cases (N=12). Zhou et al.[15] used a test set consisting of IDH-mutated LGG and high-grade glioma to evaluate the 1p/19q co-deletion prediction performance. This is not an ideal test set as 1p/19q co-deletion status is not clinically relevant for high grade glioma, and there is a selection bias of IDH mutated tumors only.

The aim of this retrospective study was to develop a radiogenomics approach to predict the 1p/19q co-deletion status of presumed LGG based on pre-operative MRI features, with a machine learning algorithm that was validated on a large external dataset.

Methods

EMC/HMC Dataset

Study participants

All patients aged 18 years or older newly diagnosed with presumed LGG and who underwent tumor resection or biopsy between October 2002 and March 2017 at the Erasmus MC, University Medical Center Rotterdam (EMC) or Haaglanden Medical Center (HMC) were retrospectively included in the EMC/HMC dataset. Patients were eligible if histopathological diagnosis with molecular subclassification of the 1p/19q co-deletion status and pre-operative post-contrast T1-weighted and T2-weighted MRI scans were available. The study was approved by the Medical Ethical Committee of Erasmus MC, who waived the need for written informed consent from the patients due to the retrospective nature of this study and the (emotional) burden that would result from contacting the patients or their relatives to obtain consent. The study was performed in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Histopathological diagnosis and molecular subclassification

Tumor samples were obtained from patients who underwent surgical resection or biopsy. Histopathological examination was performed by neuropathologists and further molecular subclassification of the 1p/19q co-deletion and/or IDH mutation status was performed as part of the diagnostic routine by molecular biologists using fluorescence in situ hybridization (FISH), Loss of Heterozygosity (LOH) analysis, targeted Next-Generation Sequencing (NGS) panel using an Ion Torrent Personal Genome Machine (Life Technologies) or Ion S5XL or a Multiplex Ligation Probe Assay (MRC-Holland). [4,16-18] All tumors were subclassified based on the WHO 2016 criteria.

Imaging acquisition and post-processing

MRI scans were used that were acquired in the routine diagnostic process. T1-weighted and T2-weighted MRI sequences were used for the algorithm. In many, but not all, patients T2-weighted fluid attenuated inversion recovery (T2w-FLAIR) imaging was also available. As scans were acquired at a number of sites, the imaging data were heterogeneous with a wide range of acquisition settings in voxel spacing, matrix size, echo time, repetition time, number of slices, slice thickness, and field strengths on scanners from three different manufacturers (General Electric, Philips and Siemens). An overview of the scanning settings is given in the Supplementary Materials, Appendix 1.

All scans were visually inspected by M.S. and excluded if MRI artefacts were present. Presumed LGG was defined as non-enhancing tumor, as seen on the presurgical post-contrast T1-weighted MRI scan. Therefore, all post-contrast T1-weighted scans were reviewed and excluded if clear or solid enhancement was present. When available T1-weighted pre-contrast scans were inspected for hemorrhage, to prevent false positive

assessment of enhancement. Although tumors with evident contrast enhancement were excluded, minimal enhancement was tolerated. Minimal enhancement was defined as punctiform (<1mm in diameter) or poorly defined faint enhancement, similar to Pallud et al.[16]

Tumor segmentation was performed by two independent observers (F.I. and G.K.) using ITKSnap.[20] Segmentation was done on T2w-FLAIR when available (N=119), otherwise on the T2-weighted scans (N=165). Since in our institution LGG segmentations are preferably performed on T2w-FLAIR scans, we did not enforce the assessors to segment on T2-weighted scans in order to stick to the real-world clinical practice. The segmentations were then transformed to the T2-weighted scans (in the case of T2w-FLAIR segmentation) and the T1-weighted scans, using the image registration software SimpleElastix[21]. For all patients, brain masks were automatically constructed using FSL's BET tool with a fractional intensity threshold of 0.5.[22] These brain masks were subsequently used to normalize the intensity of the MRI scans. Details can be found in Appendix 2.

TCIA Dataset

Patients from The Cancer Imaging Archive (TCIA) "LGG-1p19qDeletion" dataset were screened for eligibility based on previously described inclusion and exclusion criteria, and used as the external validation dataset.[10,23,24]

This data collection is a publicly available dataset that consists of histopathological proven LGG with co-registered T1- and T2-weighted preoperative MRI scans as well as biopsy proven 1p/19q co-deletion status. Molecular analysis of the 1p/19q co-deletion status was performed with FISH for all tumors; IDH mutation status was not determined. All MRI scans were visually inspected by M.S. as previously described. An overview of the MRI settings is listed in the Supplementary Materials, Appendix 1. All tumors were semi-automatically segmented by M.S. on the T2-weighted scans using ITKSnap. Since the T1-weighted and T2-weighted scans were already co-registered in this study, the segmentation could be directly used for the T1-weighted scans without the need for registration. Brain masks were made using FSL's BET tool, with the same settings as for the EMC/HMC dataset.

Classification algorithm

To predict the 1p/19q status of the tumors based on MRI features, the PREDICT toolbox was used. This toolbox was used to extract a total of 78 image features (such as image intensity, tumor texture, tumor shape, and tumor location) from the T1-weighted and T2-weighted MR image. These features, as well as the age and sex of the patient, were then used to train a Support Vector Machine (SVM), resulting in a total of 80 features. All parameter optimization and classifier training were performed on the EMC/HMC training set dataset using 100 iterations of stratified random-split cross-validation, with 80% of the data set used for training and 20% used for validation. Once the algorithm was optimized, no more

changes were made to the algorithm and it was then evaluated on the TCIA dataset. To evaluate the algorithm, the accuracy, sensitivity (1p/19q co-deletion prediction), specificity (1p/19q intact prediction), area under the receiver-operating-characteristic curve (AUC), weighted F1-score, and precision were determined by comparing the predicted labels with the reference labels obtained from tissue diagnosis. Full details of the algorithm can be found in the Supplementary Materials, Appendix 2 with more information about the evaluation metrics in Appendix 3. An overview of the classification algorithm is provided in Supplementary Materials, Online Figure 1.

To minimize the variance due to randomness in the algorithm training, an ensemble of 5 SVMs, which averages the predictions of the 5 independently trained models, was also constructed; the details can be found in Supplementary Materials, Appendix 2. One hundred different ensembles were constructed, and were evaluated on the TCIA dataset using the evaluation metrics described previously. Mean and standard deviation of the metrics over the 100 ensembles were computed.

To evaluate the contribution of the different features to the final prediction, a sensitivity analysis using polynomial chaos expansions was performed, resulting in Sobol indices for each feature.[25] The total Sobol index was used to determine the relative feature importance of the individual features. The total Sobol index is relative measure of the sensitivity of the algorithm to the input features. The OpenPC toolbox was used to create the polynomial chaos expansions and to calculate the Sobol indices.[26,27]

We also determined which patients from the TCIA dataset were considered as representative examples for the 1p/19q co-deleted and 1p/19q intact class by the algorithm. This was achieved by counting the number of times the algorithm correctly predicted the class for a specific patient in the 100 ensembles that were constructed. We evaluated the performance of the algorithm when the EMC/HMC and TCIA dataset were mixed instead of used as a separate train and validation set, to evaluate the effect of adding additional training data.

Prediction of 1p/19q status by clinical experts

To compare the results of the algorithm with expert performance, the 1p/19q status of the TCIA tumors was also predicted by two neuroradiologists and two neurosurgeons at the Erasmus MC Brain Tumor Center. They were presented with the T1-weighted and T2-weighted images side by side for each patient as well as the sex and age to ensure that the algorithm and the raters had access to the same information. For each tumor the rater was then asked to choose whether they thought it was 1p/19q co-deleted or not, and to provide a confidence score ranging from 1 to 5 (1 indicating very unsure and 5 indicating very sure). This confidence score was then turned into a prediction 'score' by dividing it by 5 and multiplying it by 1 if the predicted label was co-deleted or by -1 if the predicted label was not co-deleted. In this way an AUC could be determined for the

manual classification. The accuracy, sensitivity, and specificity were determined in the same way as for the algorithm.

Statistical analysis

Statistical analyses to test differences between the two datasets were performed with SPSS 21.0 statistical software (IBM Corp.). We tested whether the two datasets differed significantly from each other using the Mann–Whitney U test for continuous, non-normally distributed variables (age and volume) and the chi-squared test for all other categorical variables (sex, genetic analysis, presence of mild enhancement, codeletion status). Predictive performances (mean, 95% confidence interval (CI)) between the EMC/HMC training set and TCIA validation set were tested with the Welch t-test. Accuracy between the clinical experts and the algorithm were tested with the McNemar test. A p-value of < 0.05 was considered statistically significant. The 95% CIs were calculated such that if the entire experiment of training on EMC/HMC and prediction on TCIA would be repeated, in 95% of the repetitions the result would lie within that interval.

Data Sharing

The data used in this study is available on Mendeley Data (http://dx.doi.org/10.17632/rssf5nxxby.1). The code for the construction and evaluation of the prediction algorithm is available on GitHub (https://github.com/Svdvoort/PREDICT). The code used to construct the polynomial chaos expansions and calculate the Sobol indices is available on GitHub as well (https://github.com/Svdvoort/OpenPC).

Results

In the EMC/HMC dataset, 424 LGG were identified and screened for eligibility. Cases were excluded due to unknown 1p/19q co-deletion status (N=22), absence of T1- and/or T2-weighted MRI scans (N=46), enhancement (N=58), and unacceptable image quality (N=14), which resulted in 284 patients included for final analysis (Flowchart, Figure 1).

From the TCIA database, all 159 patients were screened for eligibility. Patients were excluded because of enhancement (N=18), signs of prior biopsy/surgical procedure (N=7), no post-contrast T1-weighted imaging available (N=3), and patients being younger than 18 years (N=2), resulting in 129 patients included in the external validation dataset (Flowchart, Figure 1). An overview of the excluded patients from the TCIA database as well as the reason for exclusion is available as Supplementary Material Appendix 4.

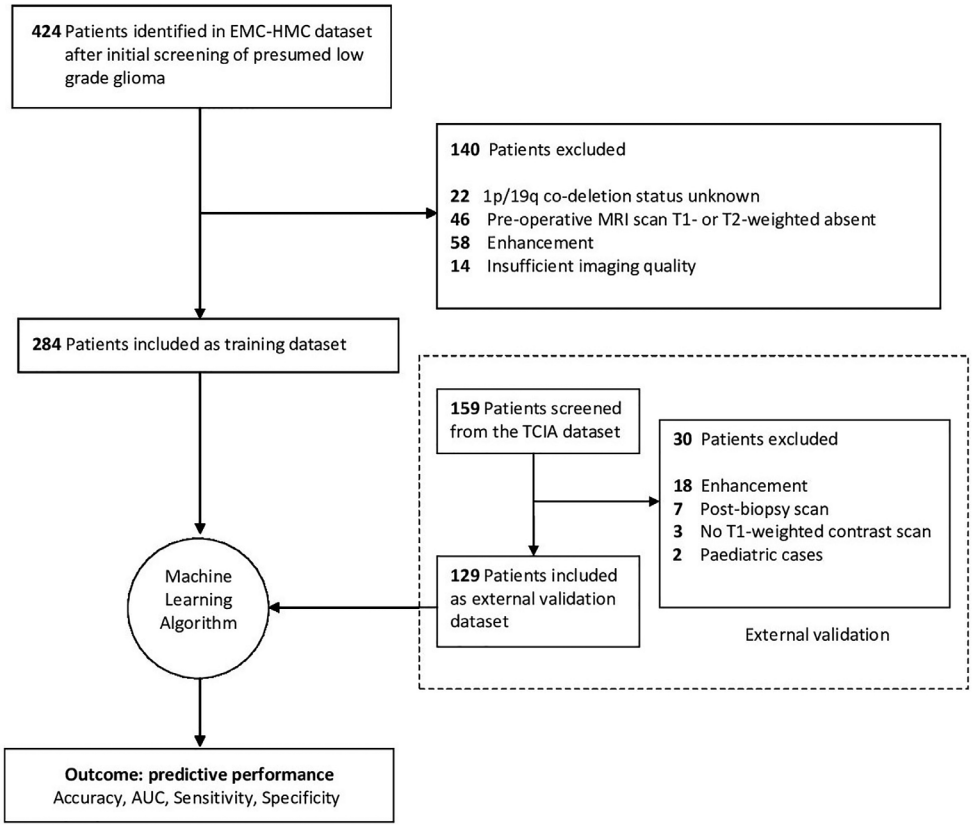


Figure 1. Flow diagram of the inclusion procedure for both the EMC/HMC training dataset and TCIA validation dataset.

There was no significant difference between the EMC/HMC and TCIA datasets for median age (43.0 years, 17.0 interquartile range (IQR) vs. 39 years, IQR 19.5 respectively; p=0.11) and sex distribution (56.7% vs 52.7% male respectively; p=045). Median tumor volume in the EMC/HMC dataset was significantly larger than in the TCIA dataset (median 47.80 cm^3, IQR 58.65 vs. median 35.70 cm^3, IQR 49.10), p=0.04). There were fewer 1p/19q co-deleted tumors in the EMC/HMC compared with the TCIA dataset (35.20% vs. 65.40%, p<0.0001). Patient and tumor characteristics of both datasets are further presented in Table 1.

The predictive performance of the algorithm on the EMC/HMC training dataset, obtained from the cross validation, and the TCIA validation dataset is given in terms of accuracy, AUC, F1-score, precision, sensitivity and specificity in Table 2. The accuracy, AUC and sensitivity did not differ significantly between training and validation datasets (p=0.886, p=0.746, p=0.146 respectively), while the specificity was significantly lower in the validation dataset (p=0.038). The predictive performances of the clinical experts compared to the algorithm can be found in Table 3, and their ROC curves in Figure 2.

The algorithm had a higher AUC when compared with the average performance of the neurosurgeons, but a lower AUC when compared with the neuroradiologists. There was high variability in predictive performance between the clinical experts (AUC 0.449 - 0.830).

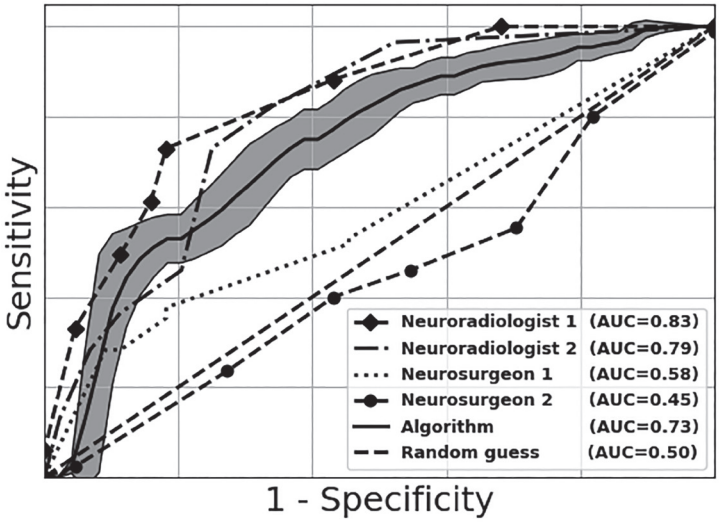


Figure 2. ROC curves of clinical expert and algorithm performance. For the performance of the algorithm the 95% confidence interval is plotted as well, representing the uncertainty due to randomness in model training.

The results of mixing the EMC/HMC dataset and the TCIA dataset are shown in Appendix 5. Mixing the datasets leads to a slightly improvement performance, but still within the confidence interval of the EMC/HMC dataset cross-validation results. According to the algorithm, the most important features for accurate 1p/19q co-deletion status prediction were the cranial/caudal location of the tumor, the skewness of the T2-weighted signal intensity (SI) histogram and one of the texture features, together with age and sex (Supplementary Materials, Online Figure 2). The algorithm identified a typical 1p/19q codeleted glioma as a frontal heterogeneous tumor as seen on T1-weighted and T2-weighted scans, while a typical 1p/19q intact glioma was identified as a parietal homogenous tumor, as shown in Figure 3.

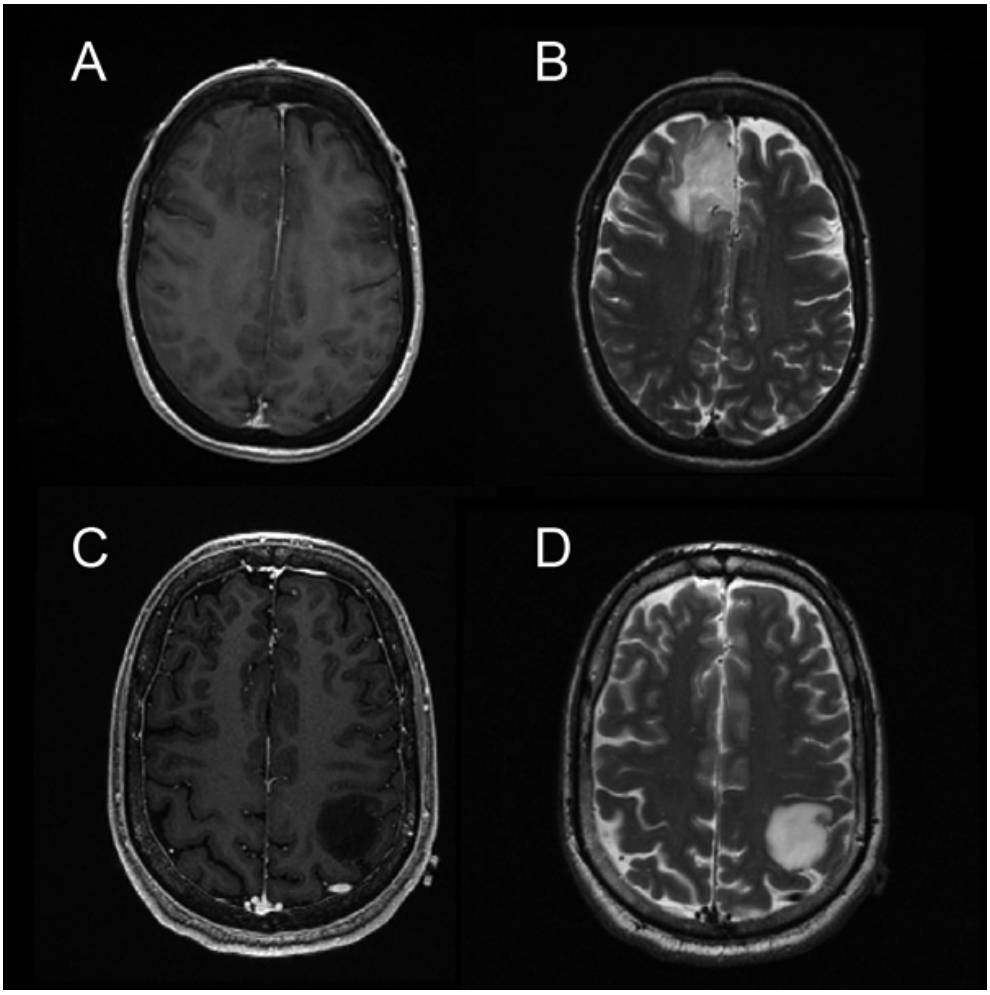


Figure 3. (A) and (B) show a frontally located glioma. It is non-enhancing on post-contrast T1-weighted MRI (A). A heterogeneous signal intensity with indistinct border is visible on the T2-weighted MRI (B). Correctly predicted as a 1p/19q codeleted glioma (oligodendroglioma) by the algorithm. (C) and (D) show a parietally located glioma. It is non-enhancing on post-contrast T1-weighted MRI (C). A homogeneous signal intensity with sharply demarcated border is visible on the T2-weighted MRI (D). Correctly predicted as a 1p/19q intact glioma (astrocytoma) by the algorithm.

Table 1. Patient and tumor characteristics.

	EMC/HMC - Training Set (n = 284)	TCIA - Validation Set (n = 129)	P value
Clinical	n (%)	n (%)	
Age median, [IQR] in years	43 [17]	39 [19.5]	0.11
Sex			0.45
Male	161 (56.7)	68 (52.7)	
Female	123 (43.3)	61 (47.3)	
Imaging			
Volume median, [IQR] in cm^3	47.8, [58.7]	35.7, [49.1]	0.04
Mild Enhancement			0.005
Yes	27 (9.5)	25 (19.4)	
No	257 (90.5)	104 (80.6)	
Genetic			
Histopathology (WHO 2016)			< 0.0001
Oligodendroglioma	100	85	
Astrocytoma	181	44	
Glioblastoma	3	0	
Method of Analysis			< 0.0001
NGS	214 (75.4)	0 (0)	
FISH	45 (15.8)	129 (100)	
MLPA	25 (8.8)	0 (0)	
1p/19q codeletion			<0.0001
Yes	100 (35.2)	85 (65.9)	
No	184 (64.8)	44 (34.1)	
IDH mutation			n/a
Yes	214 (75.4)	0 (0.0)	
No	35 (12.3)	0 (0.0)	
Unknown	35 (12.3)	129 (100.0)	

EMC Erasmus MC, HMC Haaglanden Medical Center, TCIA The Cancer Imaging Archive, NGS Next Generation Sequencing, FISH Fluorescent in situ hybridization, MLPA Multiplex Ligation Probe Assay, WHO World Health Organization.

Table 2. Predictive performances of the algorithm on the EMC/HMC training and TCIA validation datasets. The performances on the EMC/HMC training dataset were obtained by cross-validation, the performances on the TCIA validation dataset were obtained by training on the EMC/HMC dataset and then testing on the TCIA dataset.

	EMC/HMC - Training Set Mean (95% CI)	TCIA - Validation Set Mean (95% CI)	P value
Accuracy	0.698 (0.636 - 0.760)	0.693 (0.657 - 0.729)	0.872
AUC	0.755 (0.694 - 0.817)	0.723 (0.708 - 0.737)	0.313
F1-score	0.701 (0.640 - 0.761)	0.697 (0.661 - 0.733)	0.896
Precision	0.570 (0.491 - 0.649)	0.787 (0.754 - 0.820)	< 0.001
Sensitivity	0.657 (0.562 - 0.752)	0.732 (0.689 - 0.775)	0.123
Specificity	0.721 (0.628 - 0.813)	0.617 (0.544 - 0.691)	0.027

AUC area under the curve.

Table 3. Predictive performance of four clinical experts compared with the algorithm on the TCIA Validation Dataset.

	Neuro-surgeon 1	Neuro-surgeon 2	Average of surgeons	Neuro-radiologist 1	Neuro-radiologist 2	Average of radiologists	Algorithm
Accuracy, with p value*	0.520, 0.073	0.457, 0.002	0.489	0.690, 0.720	0.574, 0.266	0.632	0.693
AUC	0.580	0.449	0.515	0.830	0.792	0.811	0.723
Sensitivity	0.370	0.459	0.415	0.610	0.459	0.535	0.732
Specificity	0.820	0.455	0.638	0.840	0.795	0.818	0.617

* statistical comparison (McNemar) of accuracy between clinical experts and algorithm. AUC area under the curve.

Discussion

In this study, we developed an algorithm that predicted the 1p/19q co-deletion status of presumed LGG non-invasively based on preoperative MRI scans with an AUC of approximately 0.75. We tested the algorithm on an external, independent validation dataset. To the best of our knowledge, this is the first time this has been done in presumed LGG, and thus sets a benchmark for the expected performance in the real-world clinical setting. The algorithm had a higher AUC than the averaged AUC of the neurosurgeons, but lower than the averaged AUC of the neuroradiologists.

To the best of our knowledge, this is the first study performing a radiogenomics based machine learning study in LGG from the perspective of real-world clinical practice: we included all patients with presumed, non-contrast enhancing LGG, rather than a selection of histopathological defined LGG patients. This is important, since in a clinical setting the genetic mutation is unknown at first symptomatic presentation. Since it is only known after surgery and molecular analysis, we aimed to mirror this real-world situation as best as possible by not selecting patients based on histological tumor features, but on the imaging features that are available at the time of presentation. Note that subsequently all lesions were surgically resected to obtain the ground truth data based on confirmed

histological and molecular analysis. We trained the algorithm on a heterogeneous training dataset and used a separate, completely independent, publicly available dataset with data from an entirely different institute to validate the algorithm. As such, this study is the first to demonstrate that the performance of a radiogenomics algorithm in predicting the 1p/19q co-deletion status of presumed LGG based on MRI scans was robust and matched expert clinical performance. Furthermore, we were also able to show which image features were important in the classification, increasing the clinical understanding of the machine learning algorithm and potentially aiding better acceptance, as well as furthering fundamental research into understanding of glioma pathophysiology.

Although other studies did already investigate the non-invasive prediction of the molecular subtype of LGG, these often focused on IDH mutations only and did not consider the 1p/19q co-deletion status.[11,28,29] In comparison with studies that did look at the 1p/19q co-deletion, we used a larger cohort and an external validation dataset,[10,13,14,15,30,31] which makes our results more robust and generalizable respectively. Although one study did use an independent dataset, this study used only five patients to externally validate the 1p/19q co-deletion predictive performance of the algorithm, which severely limits the reliability of its predictive performance.[12] Additionally, that specific study retrospectively selected patients with histopathological defined LGG only, which represents the diagnosis-treatment workflow in clinical practice less accurately. The starting point of decision making on the optimal treatment strategy for LGG is the initial diagnosis on first MRI, when a non-contrast enhancing space occupying lesion is seen, at which point knowledge on the histopathological grade is not yet available.

The optimal timing and effect of surgical treatment of LGG is extensively being debated within literature and has recently been re-evaluated in the light of molecular subclassification after the introduction of WHO 2016 criteria.[5,6,32,33] Currently, the molecular subtype based on 1p/19q codeletion and IDH mutation can be diagnosed only after obtaining tissue with biopsy or surgery. Indeed, as our results suggests, it is even for experienced neuro-oncological surgeons and radiologists a challenge to accurately predict the codeletion status of non-enhancing tumors based on preoperative MRI scans (AUC 0.45-0.83).

There are two scenarios in which preoperative, non-invasive prediction of the 1p/19q codeletion status based on MRI would be clinically relevant. First, some patients are not eligible for surgical resection or diagnostic biopsy due to older age, poor neurological condition, or tumor localization in eloquent brain areas or basal ganglia.[33] However, knowledge of the molecular LGG subtype might add to a more appropriate (timing of) chemo- and/or radiotherapy regimes (immediate post-operative therapy vs watchful waiting).[34] Therefore, non-invasive, accurate prediction of the molecular subtype on imaging could help clinicians to select the optimal treatment when tissue diagnosis is difficult to obtain. Second, it is suggested that post-surgical residual 1p/19q intact, IDH mutated tumor has a more negative impact on survival than residual 1p/19q co-deleted,

IDH mutated (oligodendroglioma) tumor.[5,6] With pre-surgical knowledge of the specific molecular subtype the surgeon can make a better informed decision on whether or not to push the limits of resection at the time of surgery, avoiding on the one hand re-resection in case of residual 1p/19q intact, IDH mutated tumor and less-justified post-surgical deficits in 1p/19q co-deleted tumor on the other hand. Clearly, the diagnostic accuracy of our algorithm is as yet too low to rely on for clinical practice. However, the results are promising because they generalize through multiple datasets, encouraging future research in this direction.

Our study had a few limitations. First, for the current study only the T1-weighted and T2-weighted images were used, while diffusion weighted and perfusion imaging also contain relevant features for the 1p/19q status. These sequences were not included in the development of the present algorithm, as these were scarcely available in both datasets.

Second, the IDH mutation status was undetermined in all of the TCIA cases and in 35 cases of the EMC/HMC dataset. Since molecular subclassification according to the WHO 2016 guidelines is based on both the 1p/19q co-deletion status and IDH mutation status, it is important to predict both. Therefore, for our future work, we are expanding our database with more patients in whom the tumor IDH status is known, to eventually be able to predict all clinically relevant subtypes of presumed LGG. There was an imbalance between the EMC/HMC dataset and the TCIA dataset in terms of the number of co-deleted and intact cases. Despite this imbalance, our algorithm still shows similar performance between the cross-validation result of the EMC/HMC dataset and the performance on the TCIA test dataset.

Conclusion

Our results suggest that our algorithm can non-invasively predict the 1p/19q co-deletion status of presumed LGG with a performance that in general outperforms oncological neurosurgeons. We evaluated our algorithm on an independent, multicenter dataset, which demonstrated that our algorithm is robust and generalizable. The prediction of the 1p/19q co-deletion status by our algorithm can eventually add value to clinical decision making by tailoring the treatment strategy for patients with presumed LGG even prior to surgery.

References

1. 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**(6): 803-20.
2. van den Bent MJ, Smits M, Kros JM, Chang SM. Diffuse Infiltrating Oligodendroglioma and Astrocytoma. *J Clin Oncol* 2017; **35**(21): 2394-401.
3. Cairncross JG, Wang M, Jenkins RB, et al. Benefit from procarbazine, lomustine, and vincristine in oligodendroglial tumors is associated with mutation of IDH. *J Clin Oncol* 2014; **32**(8): 783-90.
4. Dubbink HJ, Atmodimedjo PN, Kros JM, et al. Molecular classification of anaplastic oligodendroglioma using next-generation sequencing: a report of the prospective randomized EORTC Brain Tumor Group 26951 phase III trial. *Neuro Oncol* 2016; **18**(3): 388-400.
5. Wijnenga MMJ, French PJ, Dubbink HJ, et al. The impact of surgery in molecularly defined low-grade glioma: an integrated clinical, radiological, and molecular analysis. *Neuro Oncol* 2018; **20**(1): 103-12.
6. Clark VE, Cahill DP. Extent of Resection Versus Molecular Classification: What Matters When? *Neurosurg Clin N Am* 2019; **30**(1): 95-101.
7. Smits M, van den Bent MJ. Imaging Correlates of Adult Glioma Genotypes. *Radiology* 2017; **284**(2): 316-31.
8. Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. *Radiology* 2014; **273**(1): 168-74.
9. Gutman DA, Cooper LA, Hwang SN, et al. MR imaging predictors of molecular profile and survival: multi-institutional study of the TCGA glioblastoma data set. *Radiology* 2013; **267**(2): 560-9.
10. Akkus Z, Ali I, Sedlar J, et al. Predicting Deletion of Chromosomal Arms 1p/19q in Low-Grade Gliomas from MR Images Using Machine Intelligence. *J Digit Imaging* 2017; **30**(4): 469-76.
11. Li Z, Wang Y, Yu J, Guo Y, Cao W. Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci Rep* 2017; **7**(1): 5467.
12. Lu CF, Hsu FT, Hsieh KL, et al. Machine Learning-Based Radiomics for Molecular Subtyping of Gliomas. *Clin Cancer Res* 2018; **24**(18): 4429-36.
13. Chang P, Grinband J, Weinberg BD et al. Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. *American Journal of Neuroradiology*. 2018;**39**(7):1201-7.
14. Han Y, Xie Z, Zang Y et al. Non-invasive genotype prediction of chromosome 1p/19q co-deletion by development and validation of an MRI-based radiomics signature in lower-grade gliomas. *Journal of neuro-oncology*. 2018;**140**(2):297-306.
15. Zhou H, Chang K, Bai HX et al. Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low-and high-grade gliomas. *Journal of neuro-oncology*. 2019;**142**(2):299-307.
16. Dubbink HJ, Atmodimedjo PN, van Marion R, et al. Diagnostic Detection of Allelic Losses and Imbalances by Next-Generation Sequencing: 1p/19q Co-Deletion Analysis of Gliomas. *J Mol Diagn* 2016; **18**(5): 775-86.

17. Riemenschneider MJ, Jeuken JW, Wesseling P, Reifenberger G. Molecular diagnostics of gliomas: state of the art. *Acta Neuropathol* 2010; **120**(5): 567-84.
18. Bieńkowski M, Wöhrer A, Moser P, Kitzwögerer M, Ricken G, Ströbel T et al. Molecular diagnostic testing of diffuse gliomas in the real-life setting: A practical approach. Clin Neuropathol. 2018 Jul-Aug; **37**(4): 166–177.
19. Pallud J, Capelle L, Taillandier L, et al. Prognostic significance of imaging contrast enhancement for WHO grade II gliomas. *Neuro Oncol* 2009; **11**(2): 176-82.
20. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* 2006; **31**(3): 1116-28.
21. Marstal K, Berendsen F, Staring M, Klein S. SimpleElastix: A user-friendly, multi-lingual library for medical image registration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Workshops* 2016: 134-42.
22. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp* 2002; **17**(3): 143-55.
23. Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 2013; **26**(6): 1045-57.
24. Erickson B, Akkus Z, Sedlar J, Kofiatis P. Data From LGG-1p19qDeletion. *The Cancer Imaging Archive* 2017. DOI: 10.7937/K9/TCIA.2017.dwehtz9v
25. Sudret B. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety* 2008; **93**(7): 964-79.
26. Perko Z, van der Voort SR, van de Water S, Hartman CM, Hoogeman M, Lathouwers D. Fast and accurate sensitivity analysis of IMPT treatment plans using Polynomial Chaos Expansion. *Phys Med Biol* 2016; **61**(12): 4646-64.
27. van der Voort S, van de Water S, Perko Z, Heijmen B, Lathouwers D, Hoogeman M. Robustness Recipes for Minimax Robust Optimization in Intensity Modulated Proton Therapy for Oropharyngeal Cancer Patients. *Int J Radiat Oncol Biol Phys* 2016; **95**(1): 163-70.
28. Ren Y, Zhang X, Rui W, et al. Noninvasive Prediction of IDH1 Mutation and ATRX Expression Loss in Low-Grade Gliomas Using Multiparametric MR Radiomic Features. *J Magn Reson Imaging* 2018.
29. Yu J, Shi Z, Lian Y, et al. Noninvasive IDH1 mutation estimation based on a quantitative radiomics approach for grade II glioma. *Eur Radiol* 2017; **27**(8): 3509-22.
30. Park YW, Han K, Ahn SS, et al. Prediction of IDH1-Mutation and 1p/19q-Codeletion Status Using Preoperative MR Imaging Phenotypes in Lower Grade Gliomas. *AJNR Am J Neuroradiol* 2018; **39**(1): 37-42.
31. Shofty B, Artzi M, Ben Bashat D, et al. MRI radiomics analysis of molecular alterations in low-grade gliomas. *Int J Comput Assist Radiol Surg* 2018; **13**(4): 563-71.
32. Jakola AS, Skjulsvik AJ, Myrmel KS, et al. Surgical resection versus watchful waiting in low-grade gliomas. *Ann Oncol* 2017; **28**(8): 1942-8.
33. Jiang B, Chaichana K, Veeravagu A, Chang SD, Black KL, Patil CG. Biopsy versus resection for the management of low-grade gliomas. *Cochrane Database Syst Rev* 2017; **4**: CD009319.

34. Ricard D, Kaloshi G, Amiel-Benouaich A, Lejeune J, Marie Y, Mandonnet E et al. Dynamic history of low-grade gliomas before and after temozolomide treatment. Ann Neurol. 2007 May;61(5):484-90.

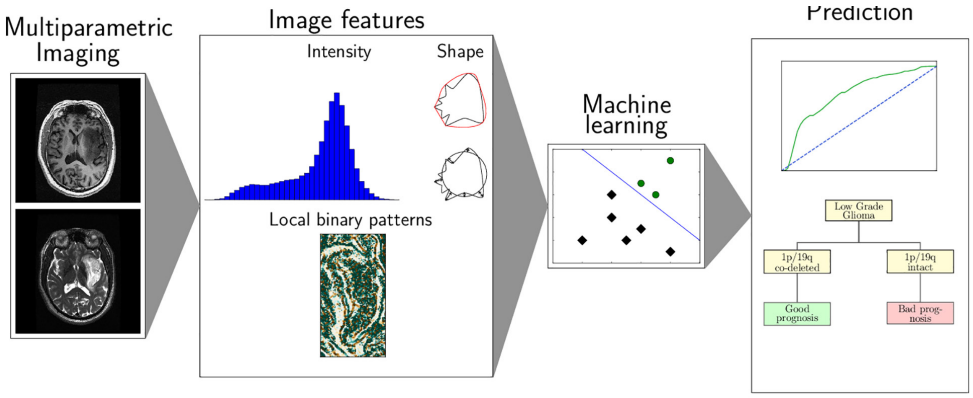
Supplementary materials

Appendix 1

Overview of MRI settings from the EMC/HMC and TCIA datasets.

	Training dataset (EMC/HMC)		Validation dataset (TCIA)	
MRI Setting	**T1-weighted min-max**	**T2-weighted min-max**	**T1-weighted min-max**	**T2-weighted min-max**
Voxel spacing in-plane (mm)	0.38 x 0.38 - 1.13 x 1.13	0.23 x 0.23 - 1.02 x 1.02	0.47 x 0.47 - 1.1 x 1.1	0.43 x 0.43 - 1.1 x 1.1
Matrix Size	256 x 176 - 1024 x 307	256 x 224 - 1024 x 1024	256 x 256 - 512x512	256 x 256 - 512x512
Echo Time (ms)	1.7 - 20	79.2 - 379	2.6 - 21	12.3 - 108.3
Repetition Time (ms)	3.8 - 1940	2000 - 13468.5	8.2 - 983.3	2033.3 - 8116.6
Slice Thickness (mm)	0.9 - 7.2	1.0 - 7.2	1.0 - 5.0	2.0 - 5.0
Number of Slices	19 - 248	19 - 304	20 - 196	20 - 84
Field Strength (Tesla)	0.5, 1.5 or 3.0	0.5, 1.5 or 3.0	1.5 or 3	1.5 or 3
# of 2D scans	147	264	33	129
# of 3D scans	137	20	96	0

Appendix 2



The radiogenomics algorithm.

An in-house radiomics pipeline, 'PREDICT', was used for the prediction of the 1p/19q co-deletion status (https://github.com/Svdvoort/PREDICT). This pipeline takes the T1- and T2-weighted MR images, tumor and brain segmentations on both scans, and labels indicating 1p/19q co-deletion status for each patient which can then be used to train a support vector machine (SVM)[1]. This SVM can then be used to predict the labels for unseen images. In this appendix the different steps of the pipeline are explained.

Data pre-processing

First, the segmentations were transformed to the T2-weighted scan and the T1-weighted scans. If the segmentation was done on T2w-FLAIR, the T2w-FLAIR scan was registered to the T2-weighted scan using SimpleElastix with an affine transform by maximisation of mutual information.[18] The computed transform was used to map the T2w-FLAIR segmentation to the T2-weighted scan. Then, the T2-weighted scans were registered to the T1-weighted scans. The computed transformation was used to transform the segmentation from the T2-weighted scans to the T1-weighted scan. This method was applied to both the segmentations originally done on the T2-weighted scans, and the ones that were registered to the T2-weighted scan from the T2w-FLAIR. All segmentations and registrations were checked by M.S., and manual adjustments of the segmentations were made by F.I. if necessary.

The next step of the pre-processing was to obtain the brain masks, which were created by FSL BET[1] with a setting of 0.5. The brain masks were used to normalize the scans. This was done by extracting the intensity values that lie within the brain mask, after which Z-scoring was applied. In this way, the mean intensity within the brain mask was 0 and the standard deviation of the intensities within the brain mask 1. These pre-processed images were then processed by the next step of the algorithm to extract the features.

Features

In total 80 features were used in our algorithm: 78 features from the T1-weighted and the T2-weighted images, and age and sex. These features were split into 5 groups: tumor intensity, tumor texture, tumor shape, tumor location, and demographic features. All of these features were only extracted within the tumor mask.

Image intensity was described using 11 features: minimum, maximum, mean, median, standard deviation, skewness, kurtosis, range from the 2nd to 98th percentile, range from the 25th to 75th percentile (the quartile range), the mode position (the bin in the intensity histogram which had the highest occurrence) of the intensity and the energy defined as:

$$E = \sum \left(\left(I + \min(I) \right)^2 \right)$$, where *I* is the image intensity

These features were calculated over the entire tumor in 3D. Rotation invariant Local Binary Patterns (LBPs) were used to describe the texture of the tumor[2]. The advantage of LBPs is that they are grayscale invariant, meaning that not the intensity values themselves, but the differences between intensity values of different voxels are taken into account. This is an advantage when using MRI scans, because in weighted MRI scans the absolute values signal intensities themselves do not carry information, it is only the difference in gray values in the image. LBP features were determined for a radius of 1, 3 and 5 with 8, 24 and 24 points, respectively. The LBP was determined for each slice of the MRI scan in

which tumor was present. The results over all pixels within the tumor and slices containing tumor were then concatenated and consequently the mean, standard deviation, median, kurtosis and skewness, and peak value were calculated for each setting, resulting in 18 LBP features. The intensity features and the texture features were calculated on both the T1-weighted and T2-weighted scan, resulting in 58 features.

Shape features were derived from the segmentation on the T2-weighted image. Shape features consisted of the compactness, radial distance, roughness, convexity, circular variance (cvar), ratio of principal axes (prax), elliptical variance (evar) and solidity [4,5]. All of these features were calculated on a per-slice basis (only for slices that contained tumor), after which the mean and standard deviation of the features over the slices were included as features in the algorithm, resulting in 16 features. The volume was also taking as a 3D shape feature, resulting in 17 shape features.

Tumor location was determined by using the center of mass (COM) from the brain mask and COM of the tumor segmentation. The anterior/posterior, left/right and inferior/superior coordinate of the vector pointing from the COM of the brain to the COM of the tumor were then included as features, resulting in 3 location features. Patient age and patient sex were included as the demographic features.

An overview of all the features is given in Table 1, along with the Imaging Biomarker Standardization Initiative (IBSI) code when the feature matched with one included in the IBSI[6]. Many of our features are based on local binary patterns, which are currently considered outside the scope of IBSI. However, preliminary experiments have indicated that including these features improves the performance of the classifier (results not shown). Z-scoring was applied to the features (normalized such that the mean of the features was 0 and the standard deviation was 1). This was done based only on the training set; the resulting normalisation settings were saved to apply the same normalisation to the validation set.

Appendix 3.

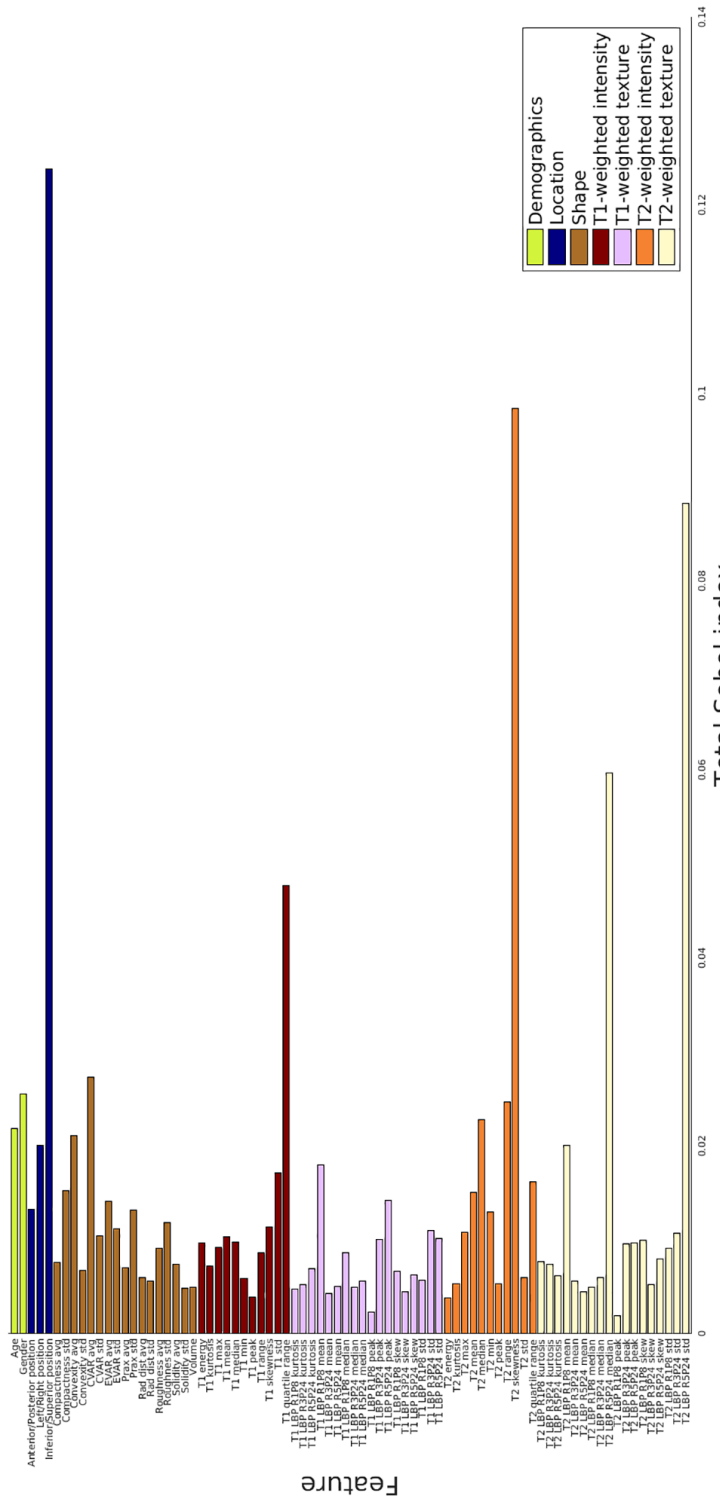

Feature
Total Sobol index
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
Demographics
Location
Shape
T1-weighted intensity
T1-weighted texture
T2-weighted intensity
T2-weighted texture
Age
Gender
Anterior/Posterior position
Left/Right position
Inferior/Superior position
Compactness std
Convexity std
Convexity avg
CVAR avg
CVAR std
EVAR avg
EVAR std
Prax avg
Prax std
Rad dist avg
Rad dist std
Roughness avg
Roughness std
Solidity avg
Solidity std
Volume
T1 energy
T1 kurtosis
T1 max
T1 mean
T1 median
T1 min
T1 peak
T1 range
T1 skewness
T1 std
T1 quantile range
T1 LBP R1P8 kurtosis
T1 LBP R3P24 kurtosis
T1 LBP R1P8 mean
T1 LBP R3P24 mean
T1 LBP R1P8 median
T1 LBP R3P24 median
T1 LBP R1P8 peak
T1 LBP R3P24 peak
T1 LBP R1P8 skew
T1 LBP R3P24 skew
T1 LBP R1P8 std
T1 LBP R3P24 std
T2 energy
T2 kurtosis
T2 max
T2 mean
T2 median
T2 peak
T2 range
T2 skewness
T2 std
T2 quantile range
T2 LBP R1P8 kurtosis
T2 LBP R3P24 kurtosis
T2 LBP R1P8 mean
T2 LBP R3P24 mean
T2 LBP R1P8 median
T2 LBP R3P24 median
T2 LBP R1P8 peak
T2 LBP R3P24 peak
T2 LBP R1P8 skew
T2 LBP R3P24 skew
T2 LBP R1P8 std
T2 LBP R3P24 std

Table 1. Overview of the features included in the algorithm.

Modality	Slice averaging	Feature	IBSI code
N/A	N/A	Age	N/A
N/A	N/A	Gender	N/A
N/A	N/A	Anterior/Posterior position	None
N/A	N/A	Left/Right position	None
N/A	N/A	Inferior/Superior Position	None
Segmentation	Average	Compactness	KRCK
Segmentation	Standard deviation	Compactness	KRCK
Segmentation	Average	Convexity	None
Segmentation	Standard deviation	Convexity	None
Segmentation	Average	Circular variance	None
Segmentation	Standard deviation	Circular variance	None
Segmentation	Average	Elliptic variance	None
Segmentation	Standard deviation	Elliptic variance	None
Segmentation	Average	Ratio of principal axis	None
Segmentation	Standard deviation	Ratio of principal axis	None
Segmentation	Average	Radial distance	None
Segmentation	Standard deviation	Radial distance	None
Segmentation	Average	Roughness	None
Segmentation	Standard deviation	Roughness	None
Segmentation	Average	Solidity	Inverse of 7T7F
Segmentation	Standard deviation	Solidity	Inverse of 7T7F
Segmentation	None	Volume	2PR5
T1	None	Energy	Similar to N8CA, except min intensity is added
T1	None	Kurtosis	IPH6
T1	None	Max	84IY
T1	None	Mean	Q4LE
T1	None	Median	Y12H
T1	None	Min	1GSF
T1	None	Peak	None
T1	None	Range	Similar to 20JQ, except based on 2nd and 98th percentile
T1	None	Skewness	KE2A
T1	None	Standard deviation	Square root of ECT3
T1	None	Quartile range	SALO
T1	None	LBP R1P8 Kurtosis	None
T1	None	LBP R3P24 Kurtosis	None

T1	None	LBP R5P24 Kurtosis	None
T1	None	LBP R1P8 Mean	None
T1	None	LBP R3P24 Mean	None
T1	None	LBP R5P24 Mean	None
T1	None	LBP R1P8 Median	None
T1	None	LBP R3P24 Median	None
T1	None	LBP R5P24 Median	None
T1	None	LBP R1P8 Peak	None
T1	None	LBP R3P24 Peak	None
T1	None	LBP R5P24 Peak	None
T1	None	LBP R1P8 Skewness	None
T1	None	LBP R3P24 Skewness	None
T1	None	LBP R5P24 Skewness	None
T1	None	LBP R1P8 Standard deviation	None
T1	None	LBP R3P24 Standard deviation	None
T1	None	LBP R5P24 Standard deviation	None
T2	None	Energy	Similar to N8CA, except min intensity is added
T2	None	Kurtosis	IPH6
T2	None	Max	84IY
T2	None	Mean	Q4LE
T2	None	Median	Y12H
T2	None	Min	1GSF
T2	None	Peak	None
T2	None	Range	Similar to 20JQ, except based on 2nd and 98th percentile
T2	None	Skewness	KE2A
T2	None	Standard deviation	Square root of ECT3
T2	None	Quartile range	SALO
T2	None	LBP R1P8 Kurtosis	None
T2	None	LBP R3P24 Kurtosis	None
T2	None	LBP R5P24 Kurtosis	None
T2	None	LBP R1P8 Mean	None
T2	None	LBP R3P24 Mean	None
T2	None	LBP R5P24 Mean	None
T2	None	LBP R1P8 Median	None
T2	None	LBP R3P24 Median	None
T2	None	LBP R5P24 Median	None
T2	None	LBP R1P8 Peak	None
T2	None	LBP R3P24 Peak	None

T2	None	LBP R5P24 Peak	None
T2	None	LBP R1P8 Skewness	None
T2	None	LBP R3P24 Skewness	None
T2	None	LBP R5P24 Skewness	None
T2	None	LBP R1P8 Standard deviation	None
T2	None	LBP R3P24 Standard deviation	None
T2	None	LBP R5P24 Standard deviation	None

The IBSI column indicates the code of the feature if it is present in the IBSI. Abbreviations: IBSI: Imaging Biomarker Standardization Initiative, LBP: local binary pattern.

Over-sampling of the minority class

There was an imbalance in the training dataset as there were more examples of 1p/19q intact than of 1p/19q co-deleted tumors. As a result, it was more difficult for the algorithm to correctly predict 1p/19q co-deleted tumors within the training dataset. To (partially) solve this problem we used Synthetic Minority Over-sampling Technique (SMOTE). SMOTE allows for the oversampling of the minority class and synthetically increased the number of samples for the 1p/19q co-deleted class[7]. Here, an SVM with a polynomial kernel was used to create the synthetic examples, with a ratio of 1 (meaning that there will be an equal number of samples from both classes) using 5 neighbours.

Classification

Support Vector Machines (SVMs) were used as classification algorithm[7]. SVMs were constructed using scikit-learn[8] with a polynomial kernel. The polynomial kernel was defined as:

$$K(X,Y) = (\gamma < X,Y > + C_0)^P$$

SVMs were constructed with a maximum of $1 \cdot 10^7$ iterations. The hyperparameters of the SVM were optimized using a 5-fold cross-validation where 20% of the training dataset was used as an internal validation set. A random search of 50,000 iterations was used to find the optimal hyperparameters. This optimization search for the optimal *C*, the regularization parameter of the SVM, as well as three parameters of the polynomial kernel: γ, which defines how much each sample is weighted, C_0, a trade-off between high-order and low-order terms, and *P*, the order of the polynomial kernel. The distributions for the parameters are presented in Table 1.

Table 2. Overview of the range of hyperparameters used in the optimization of the SVM.

Parameter	Lower bound	Upper bound
C	0	$1 \cdot 10^{6}$
P	1	7
C_0	0	$1 \cdot 10^{3}$
	$1 \cdot 10^{-5}$	1

SD standard deviation, IQR Inter Quartile Range, CE contrast enhancement, FLAIR Fluid-attenuated inversion recovery.

The optimal hyperparameters were chosen based on the settings that gave the highest average area under curve over the 5 folds. Using these optimal hyperparameters, the final SVM was then constructed using the complete training dataset.

Ensemble SVM

To increase the predictive performance and minimize the variability of the predictions, an ensemble of SVMs was constructed. Five trained SVMs were taken to form the ensemble. All 5 SVMs were then used to predict a test sample. The posteriors of the SVMs were averaged for the sample. Based on this averaged posterior the final label was determined (positive class if posterior >=0.5, negative class if posterior <0.5). To form the 100 ensembles 500 SVMS had to be trained. No single SVM was used in two different ensembles: all ensembles were completely unique. The resulting 100 ensemble SVMs could then be used to make a prediction of the 1p/19q status for a new sample.

Appendix 3

Evaluation metrics

The accuracy, area under the Receiver Operating Characteristic (ROC) curve (AUC), weighted F1-score, precision, sensitivity, and specificity were calculated to evaluate the performance of the algorithm and clinical experts. These were based on the variables extracted from the confusion matrix:

- True Positive (TP): number of true 1p/19q co-deleted predicted as 1p/19q co-deleted
- False Positive (FP): number of true 1p/19q intact predicted as 1p/19q co-deleted
- True Negative (TN): number of true 1p/19q intact predicted as 1p/19q intact
- False Negative (FN): number of true 1p/19q co-deleted predicted as 1p/19q intact

The evaluation metrics were then defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{F1-score} = \frac{TP^2 + TP \cdot FN + FP \cdot TN + TN^2}{TP + FN + FP + TN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

The ROC curve and AUC were obtained using scikit-learn.

Appendix 4

Table of excluded patients from TCIA-LGG dataset.

Patient ID	Reason for Exclusion
LGG-223	Enhancement
LGG-234	Enhancement
LGG-241	Enhancement
LGG-254	Post-biopsy
LGG-260	Enhancement
LGG-282	Enhancement
LGG-295	Enhancement
LGG-296	Post-biopsy
LGG-307	Enhancement
LGG-310	Post-biopsy
LGG-313	Post-biopsy
LGG-334	Post-biopsy
LGG-338	Enhancement
LGG-354	Enhancement
LGG-365	Enhancement
LGG-367	Post-biopsy
LGG-377	Enhancement
LGG-387	Enhancement
LGG-500	Age < 18 years
LGG-506	Enhancement
LGG-532	Age < 18 years
LGG-545	No post-contrast T1-weighted image

LGG-558	No post-contrast T1-weighted image
LGG-561	Enhancement
LGG-563	No post-contrast T1-weighted image
LGG-594	Enhancement
LGG-600	Enhancement
LGG-601	Post-biopsy
LGG-634	Enhancement
LGG-642	Enhancement

Appendix 5

Results of mixing EMC/HMC and TCIA datasets.

We evaluated the effect of mixing the EMC/HMC and TCIA datasets, instead of using them as separate training and validation set. This was done using the same cross-validation approach as was used for the EMC/HMC dataset cross-validation. In this case all 413 patients were pooled together. Then 100 iterations of stratified random-split cross-validation with 80% of the dataset used for training and 20% of the dataset used for validation was performed. The results are shown in Table 1.

Table 1. Predictive performances of the algorithm on the mixed EMC/HMC and TCIA dataset.

	Cross-validation on combined EMC/HMC and TCIA dataset	Cross-validation on EMC/HMC dataset
	Mean (95% CI)	**Mean (95% CI)**
Accuracy	0.717 (0.670 – 0.764)	0.698 (0.636 - 0.760)
AUC	0.780 (0.729 – 0.830)	0.755 (0.694 - 0.817)
F1-score	0.717 (0.670 – 0.763)	0.701 (0.640 - 0.761)
Precision	0.676 (0.620 – 0.732)	0.570 (0.491 - 0.649)
Sensitivity	0.710 (0.637 – 0.783)	0.657 (0.562 - 0.752)
Specificity	0.723 (0.658 – 0.788)	0.721 (0.628 - 0.813)

The results for the combined EMC/HMC and TCIA dataset show a slight improvement over the cross-validation on the EMC/HMC dataset alone but they are still within the confidence interval of the EMC/HMC dataset results, except for the precision. This shows that although there is a slight improvement, our algorithm is already quite robust and adding more data (from a different source) does not lead to a large increase in performance.

References

1. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.
2. Smith, Stephen M. "Fast robust automated brain extraction." *Human Brain Mapping* 17.3 (2002): 143-155.
3. Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.7 (2002): 971-987.
4. Peura, Markus, and Jukka Iivarinen. "Efficiency of simple shape descriptors." *Aspects of Visual Form* (1997): 443-451.
5. Xu, J., Faruque, J., Beaulieu, C. F., Rubin, D., & Napel, S. (2012). A comprehensive descriptor of shape: method and application to content-based retrieval of similar appearing lesions in medical images. Journal of digital imaging, 25(1), 121-128.
6. Zwanenburg, A., Leger, S., Vallières, M., & Löck, S. (2016). Image biomarker standardisation initiative. arXiv preprint arXiv:1612.07003.
7. Chawla, Nitesh V., et al. "SMOTE: synthetic minority over-sampling technique." *Journal of Artificial Intelligence Research* 16 (2002): 321-357.
8. Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." *Machine Learning* 20.3 (1995): 273-297.
9. Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." *Journal of Machine Learning Research* 12.Oct (2011): 2825-2830.

Chapter 4

WHO 2016 subtyping and automated segmentation of glioma using multi-task deep learning

Sebastian R. van der Voort*, Fatih Incekara* , Maarten MJ Wijnenga, Georgios Kapsas, Renske Gahrmann , Joost W Schouten , Rishi Nandoe Tewarie, Geert J Lycklama, Philip C De Witt Hamer, Roelant S Eijgelaar, Pim J French, Hendrikus J Dubbink, Arnaud JPE Vincent, Wiro J Niessen, Martin J van der Bent, Marion Smits*, and Stefan Klein*

Submitted

* contributed equally

Abstract

Background

Accurate characterization of glioma is crucial for clinical decision making. A delineation of the tumor is also desirable in the initial decision stages but is a time-consuming task.

Methods

Leveraging the latest GPU capabilities, we developed a single multi-task convolutional neural network that uses the full 3D, structural, pre-operative MRI scans to can predict the IDH mutation status, the 1p/19q co-deletion status, and the grade of a tumor, while simultaneously segmenting the tumor. We trained our method using the largest, most diverse patient cohort to date containing 1508 glioma patients from 16 institutes.

Results

We tested our method on an independent dataset of 240 patients from 13 different institutes, and achieved an IDH-AUC of 0.90, 1p/19q-AUC of 0.85, grade-AUC of 0.81, and a mean whole tumor DICE score of 0.84.

Conclusion

Our method non-invasively predicts multiple, clinically relevant parameters and generalizes well to the broader clinical population.

Introduction

Glioma is the most common primary brain tumor and is one of the deadliest forms of cancer [1]. Differences in survival and treatment response of glioma are attributed to their genetic and histological features, specifically the isocitrate dehydrogenase (IDH) mutation status, the 1p/19q co-deletion status and the tumor grade [2,3]. Therefore, in 2016 the World Health Organization (WHO) updated its brain tumor classification, categorizing glioma based on these genetic and histological features [4]. In current clinical practice, these features are determined from tumor tissue. While this is not an issue in patients in whom the tumor can be resected, this is problematic when resection cannot safely be performed. In these instances, surgical biopsy is performed with the sole purpose of obtaining tissue for diagnosis, which, although relatively safe, is not without risk [5,6]. Therefore, there has been an increasing interest in complementary non-invasive alternatives that can provide the genetic and histological information used in the WHO 2016 categorization [7,8].

Magnetic resonance imaging (MRI) has been proposed as a possible candidate because of its non-invasive nature and its current place in routine clinical care [9]. Research has shown that certain MRI features, such as the tumor heterogeneity, correlate with the genetic and histological features of glioma [10,11]. This notion has popularized, in addition to already popular applications such as tumor segmentation, the use of machine learning methods for the prediction of genetic and histological features, known as radiomics [12,13,14]. Although a plethora of such methods now exist, they have found little translation to the clinic [12].

An often discussed challenge for the adoption of machine learning methods in clinical practice is the lack of standardization, resulting in heterogeneity of patient populations, imaging protocols, and scan quality [15,16]. Since machine learning methods are prone to overfitting, this heterogeneity questions the validity of such methods in a broader patient population [16]. Furthermore, it has 2 been noted that most current research concerns narrow task-specific methods that lack the context between different related tasks, which might restrict the performance of these methods [17].

An important technical limitation when using deep learning methods is the limited GPU memory, which restricts the size of models that can be trained [18]. This is a problem especially for clinical data, which is often 3D, requiring even more memory than the commonly used 2D networks. This further limits the size of these models resulting in shallower models, and the use of patches of a scan instead of using the full 3D scan as an input, which limits the amount of context these methods can extract from the scans.

Here, we present a new method that addresses the above problems. Our method consists of a single, multi-task convolutional neural network (CNN) that can predict the IDH mutation status, the 1p/19q co-deletion status, and the grade (grade II/III/IV) of a tumor, while also simultaneously segmenting the tumor, see Figure 1. To the best of our knowledge, this is the first method that provides all of this information at the same

time, allowing clinical experts to derive the WHO category from the individually predicted genetic and histological features, while also allowing them to consider or disregard specific predictions as they deem fit. Exploiting the capabilities of the latest GPUs, optimizing our implementation to reduce the memory footprint, and using distributed multiGPU training, we were able to train a model that uses the full 3D scan as an input. We trained our method using the largest, most diverse patient cohort to date, with 1508 patients included from 16 different institutes. To ensure the broad applicability of our method, we used minimal inclusion criteria, only requiring the four most commonly used MRI sequences: pre- and post-contrast T1-weighted (T1w), T2-weighted (T2w), and T2-weighted fluid attenuated inversion recovery (T2w-FLAIR) [19,20]. No constraints were placed on the patients' clinical characteristics, such as the tumor grade, or the radiological characteristics of scans, such as the scan quality. In this way, our method could capture the heterogeneity that is naturally present in clinical data. We tested our method on an independent dataset of 240 patients from 13 different institutes, to evaluate the true generalizability of our method. Our results show that we can predict multiple clinical features of glioma from MRI scans in a diverse patient population.

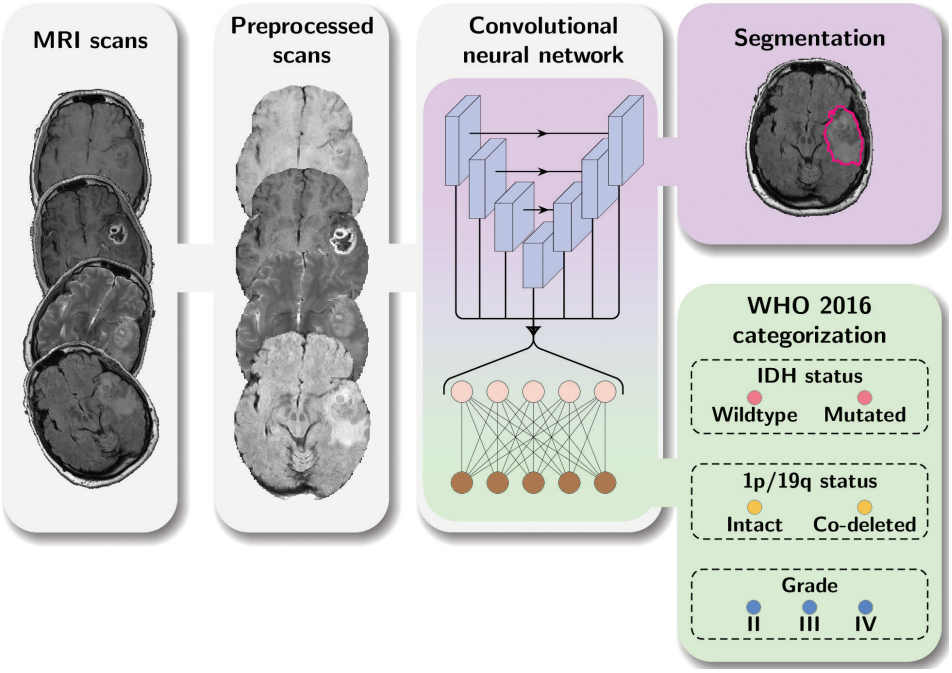


Figure 1. Overview of our method. Pre-, and post-contrast T1w, T2w and T2w- FLAIR scans are used as an input. The scans are registered to an atlas, bias field corrected, skull stripped, and normalized before being passed through our convolutional neural network. One branch of the network segments the tumor, while at the same time the features are combined to predict the IDH status, 1p/19q status, and grade of the tumor.

Methods

Patient population

The train set was collected from four in-house datasets and five publicly available datasets. In-house datasets were collected from four different institutes: Erasmus MC (EMC), Haaglanden Medical Center (HMC), Amsterdam UMC (AUMC) [37], and University Medical Center Utrecht (UMCU). Four of the five public datasets were collected from The Cancer Imaging Archive (TCIA) [38]: the Repository of Molecular Brain Neoplasia Data (REMBRANDT) collection [39], the Clinical Proteomic Tumor Analysis Consortium Glioblastoma Multiforme (CPTAC-GBM) collection [40], the Ivy Glioblastoma Atlas Project (Ivy GAP) collection [41,42], and the Brain-Tumor-Progression collection [43]. The fifth dataset was the 2019 Brain Tumor Segmentation challenge (BraTS) challenge dataset [44,45,46], from which we excluded the patients that were also available in the TCGA-LGG and TCGA-GBM collections [47,48].

For the internal datasets from the EMC and the HMC, manual segmentations were available, which were made by four different clinical experts. For patients where segmentations from more than one observer were available, we randomly picked one of the segmentations to use in the train set. The segmentations from the AUMC data were made by a single observer of the study by Visser et al. [37]. From the public datasets, only the BraTS dataset and the BrainTumor-Progression dataset provided manual segmentations. Segmentations of the BraTS dataset, as provided in the 2019 training and validation set were used. For the Brain-Tumor-Progression dataset, the segmentations as provided in the TCIA data collection were used. 15

Patients were included if pre-operative pre- and post-contrast T1w, T2w, and T2w-FLAIR scans were available; no further inclusion criteria were set. For example, patients were not excluded based on the radiological characteristics of the scan, such as low imaging quality or imaging artifacts, or the glioma's clinical characteristics such as the grade. If multiple scans of the same contrast type were available in a single scan session (e.g., multiple T2w scans), the scan upon which the segmentation was made was selected. If no segmentation was available, or the segmentation was not made based on that scan contrast, the scan with the highest axial resolution was used, where a 3D acquisition was preferred over a 2D acquisition.

For the in-house data, genetic and histological data were available for the EMC, HMC, and UMCU dataset, which were obtained from analysis of tumor tissue after biopsy or resection. Genetic and histological data of the public datasets were also available for the REMBRANDT, CPTAC-GBM, and Ivy GAP collections. Data for the REMBRANDT and CPTAC-GBM collections was collected from the clinical data available at the TCIA [39,40]. For the Ivy GAP collection, the genetic and histological data were obtained from the Swedish Institute at https://ivygap.swedish.org/home.

As a test set we used the TCGA-LGG and TCGA-GBM collections from the TCIA [47, 48]. Genetic and histological labels were obtained from the clinical data available at the TCIA. Segmentations were used as available from the TCIA, based on the 2018 BraTS challenge [45,49,50]. The inclusion criteria for the patients included in the BraTS challenge were the same as our inclusion criteria: the presence of a pre-operative pre- and post-contrast T1w, T2w, and T2w-FLAIR scan. Thus, patients from the TCGA-LGG and TCGA-GBM were included if a segmentation from the BraTS challenge was available. However, for three patients, we found that although they did have manual segmentations, they did not meet our inclusion requirements: TCGA-08-0509 and TCGA-08- 0510 from TCGA-GBM because they did not have a pre-contrast T1w scan and TCGA-FG-7634 from TCGA-LGG because there was no post-contrast T1w scan.

Automatic segmentation in the train set

To present our method with a large diversity in scans, we wanted to include as many patients in the train set as possible from the different datasets. Therefore, we performed automatic segmentation in patients that did not have manual segmentations. To this end, we used an initial version of our network (presented in Section 4.4), without the additional layers that were needed for the prediction of the genetic and histological features. This network was initially trained using all patients in the train set for whom a manual segmentation was available, and this trained network was then applied to all patients for which a manual segmentation was not available. The resulting automatic segmentations were inspected, and if their quality was acceptable, they were added to the train set. The network was then trained again, using this increased dataset, and was applied to scans that did not yet have a segmentation of acceptable quality. 16 This process was repeated until an acceptable segmentation was available for all patients, which constituted our final, complete train set.

Pre-processing

For all datasets, except for the BraTS dataset for which the scans were already provided in NIfTI format, the scans were converted from DICOM format to NIfTI format using dcm2niix version v1.0.20190410 [51]. We then registered all scans to the MNI152 T1w and T2w atlases, version ICBM 2009a, which had a resolution of 1x1x1 mm3 and a size of 197x233x189 voxels [52,53]. The scans were affinely registered using Elastix 5.0 [54,55]. The pre- and postcontrast T1w scans were registered to the T1w atlas; the T2w and T2w-FLAIR scans were registered to the T2w atlas. When a manual segmentation was available for patients from the in-house datasets, the registration parameters that resulted from registering the scan used during the segmentation were used to transform the segmentation to the atlas. In the case of the public datasets, we used the registration parameters of the T2w-FLAIR scans to transform the segmentations.

After the registration, all scans were N4 bias field corrected using SimpleITK version 1.2.4 [56]. A brain mask was made for the atlas using HD-BET, both for the T1w atlas and the T2w atlas [57]. This brain mask was used to skull strip all registered scans and crop them to a bounding box around the brain mask, reducing the amount of background present in the scans, resulting in a scan size of 152x182x145 voxels. Subsequently, the scans were normalized such that for each scan, the average image intensity was 0, and the standard deviation of the image intensity was 1 within the brain mask. Finally, the background outside the brain mask was set to the minimum intensity value within the brain mask.

Since the segmentation could sometimes be rugged at the edges after registration, especially when the segmentations were initially made on low-resolution scans, we smoothed the segmentation using a 3x3x3 median filter (this was only done in the train set). For segmentations that contained more than one label, e.g., when the tumor necrosis and enhancement were separately segmented, all labels were collapsed into a single label to obtain a single segmentation of the whole tumor. The genetic and histological labels and the segmentations of each patient were one-hot encoded. The four scans, ground truth labels, and segmentation of each patient were then used as the input to the network.

Model

We based the architecture of our model on the U-Net architecture, with some adaptations made to allow for a full 3D input and the auxiliary tasks [58]. Our network architecture, which we have named PrognosAIs Structure-Net, or PSNet for short, can be seen in Figure 8.

To use the full 3D scan as an input to the network, we replaced the first pooling layer that is usually present in the U-Net with a strided convolution, with a kernel size of 9x9x9 and a stride of 3x3x3. In the upsampling branch of the network, the last up-convolution is replaced by a deconvolution, with the same kernel size and stride.

At each depth of the network, we have added global max-pooling layers directly after the dropout layer, to obtain imaging features that can be used to predict the genetic and histological features. We chose global pooling layers as they do not introduce any additional parameters that need to be trained, thus keeping the memory required by our model manageable. The features from the different depths of the network were concatenated and fed into three different dense layers, one for each of the genetic and histological outputs.

l2 kernel regularization was used in all convolutional layers, except for the last convolutional layer used for the output of the segmentation. In total this model contained 27,042,473 trainable an 2,944 non-trainable parameters.

Model training

Training of the model was done on eight NVidia RTX2080Ti's with 11GB of memory, using TensorFlow 2.2.0 [59]. To be able to use the full 3D scan as input to the network, without

running into memory issues, we had to optimize the memory efficiency of the model training. Most importantly, we used mixedprecision training, which means that most of the variables of the model (such as the weights) were stored in float16, which requires half the memory of float32, which is typically used to store these variables [60]. Only the last softmax activation layers of each output were stored as float32. We also stored our pre-processed scans as float16 to further reduce memory usage.

However, even with these settings, we could not use a batch size larger than 1. It is known that a larger batch size is preferable, as it increases the stability of the gradient updates and allows for a better estimation of the normalization parameters in batch normalization layers [61]. Therefore, we distributed the training over the eight GPUs, using the NCCL AllReduce algorithm, which combines the gradients calculated on each GPU before calculating the update to the model parameters [62]. We also used synchronized batch normalization layers, which synchronize the updates of their parameters over the distributed models. In this way, our model had a virtual batch size of eight for the gradient updates and the batch normalization layers parameters.

To provide more samples to the algorithm and prevent potential overtraining, we applied four types of data augmentation during training: cropping, rotation, brightness shifts, and contrast shifts. Each augmentation was applied with a certain augmentation probability, which determined the probability of that augmentation type being applied to a specific sample. When an image was cropped, a random number of voxels between 0 and 20 was cropped from each dimension, and filled with zeros. For the random rotation, an angle between −30° and 30° degrees was selected from a uniform distribution for each dimension. The brightness shift was applied with a delta uniformly drawn between 0 and 0.2, and the contrast shift factor was randomly drawn between 0.85 and 1.15. We also introduced an augmentation factor, which determines how often each sample was parsed as an input sample during a single epoch, where each time it could be augmented differently.

For the IDH, 1p/19q, and grade output, we used a masked categorical crossentropy loss, and for the segmentation we used a DICE loss, see Appendix E for details. We used AdamW as an optimizer, which has shown improved generalization performance over Adam by introducing the weight decay parameter as a separate parameter from the learning rate [63]. The learning rate was automatically reduced by a factor of 0.25 if the loss did not improve during the last five epochs, with a minimum learning rate of 1 . 10−11. The model could train for a maximum of 150 epochs, and training was stopped early if the average loss over the last five epochs did not improve. Once the model was finished training, the weights from the epoch with the lowest loss were restored. 19

Hyperparameter tuning

Hyperparameters involved in the training of the model needed to be tuned to achieve the best performance. We tuned a total of six hyper parameters: the l2- norm, the dropout rate,

the augmentation factor, the augmentation probability, the optimizer's initial learning rate and the optimizer's weight decay. A full overview of the trained parameters and the values tested for the different settings is presented in Appendix F.

To tune these hyperparameters, we split the train set into a hyperparameter training set (85%/1282 patients of the full train data) and a hyperparameter validation set (15%/226 patients of the full train data). Models were trained for different hyperparameter settings via an exhaustive search using the hyperparameter train set, and then evaluated on the hyperparameter validation set. No data augmentation was applied to the hyperparameter validation to ensure that results between trained models were comparable. The hyperparameters that led to the lowest overall loss in the hyperparameter validation set were chosen as the optimal hyperparameters. We trained the final model using these optimal hyperparameters and the full train set.

Post-processing

The predictions of the network were post-processed to obtain the final predicted labels and segmentations for the samples. Since softmax activations were used for the genetic and histological outputs, a prediction between 0 and 1 was outputted for each class, where the individual predictions summed to 1. The final predicted label was then considered as the class with the highest prediction score. For the prediction of LGG (grade II/III) vs. HGG (grade IV), the prediction scores of grade II and grade III were combined to obtain the prediction score for LGG, the prediction score of grade IV was used as the prediction score for HGG. If a segmentation contained multiple unconnected components, we only retained the largest component to obtain a single whole tumor segmentation.

Model evaluation

The performance of the final trained model was evaluated on the independent test set, comparing the predicted labels with the ground truth labels. For the genetic and histological features, we evaluated the AUC, the accuracy, the sensitivity, and the specificity using scikit-learn version 0.23.1, for details see Appendix G [64]. We evaluated these metrics on the full test set and in subcategories relevant to the WHO 2016 guidelines. We evaluated the IDH performance separately in the LGG (grade II/III) and HGG (grade IV) subgroups, the 1p/19q performance in LGG, and we also evaluated the performance of distinguishing between LGG and HGG instead of predicting the individual grades.

To evaluate the performance of the segmentation, we calculated the DICE scores, Hausdorff distances, and volumetric similarity coefficient comparing the automatic segmentation of our method and the manual ground truth segmen20 tations for all patients in the test set. These metrics were calculated using the EvaluateSegmentation toolbox, version 2017.04.25 [65], for details see Appendix G.

To prevent an overly optimistic estimation of our model's predictive value, we only evaluated our model on the test set once all hyperparameters were chosen, and the final

model was trained. In this way, the performance in the test set did not influence decisions made during the development of the model, preventing possible overfitting by fine-tuning to the test set.

To gain insight into the model, we made saliency maps that show which parts of the scan contribute the most to the prediction of the CNN [66]. Saliency maps were made using tf-keras-vis 0.5.2, changing the activation function of all output layers from softmax to linear activations, using SmoothGrad to reduce the noisiness of the saliency maps [66].

Another way to gain insight into the network's behavior is to visualize the filter outputs of the convolutional layers, as they can give some idea as to what operations the network applies to the scans. We visualized the filter outputs of the last convolutional layers in the downsample and upsample path at the first depth (at an image size of 49x61x51) of our network. These filter outputs were visualized by passing a sample through the network and showing the convolutional layers' outputs, replacing the ReLU activation with linear activations.

Data availability

An overview of the patients included from the public datasets used in the training and testing of the algorithm, and their ground truth label is available in Appendix H. The data from the public datasets are available in TCIA under DOIs: 10.7937/K9/TCIA.2015.588OZUZB, 10.7937/k9/tcia.2018.3rje41q1, 10.7937/K9/TCIA.2016.XLwaN6nL, and 10.7937/K9/TCIA.2018.15quzvnb. Data from the BraTS are available at http://braintumorsegmentation.org/. Data from the in-house datasets are not publicly available due to participant privacy and consent. 4.10 Code availability The code used in this paper is available on GitHub under an Apache 2 license at https://github.com/Svdvoort/PrognosAIs_glioma. This code includes the full pipeline from registration of the patients to the final post-processing of the predictions. The trained model is also available on GitHub, along with code to apply it to new patients.

Results

Patient characteristics

We included a total of 1748 patients in our study, 1508 as a train set and 240 as an independent test set. The patients in the train set originated from nine different data collections and 16 different institutes, and the test set was collected from two different data collections and 13 different institutes. Table 1 provides a full overview of the patient characteristics in the train and test set, and Figure 2 shows the inclusion flowchart and the distribution of the patients over the different data collections in the train set and test set.

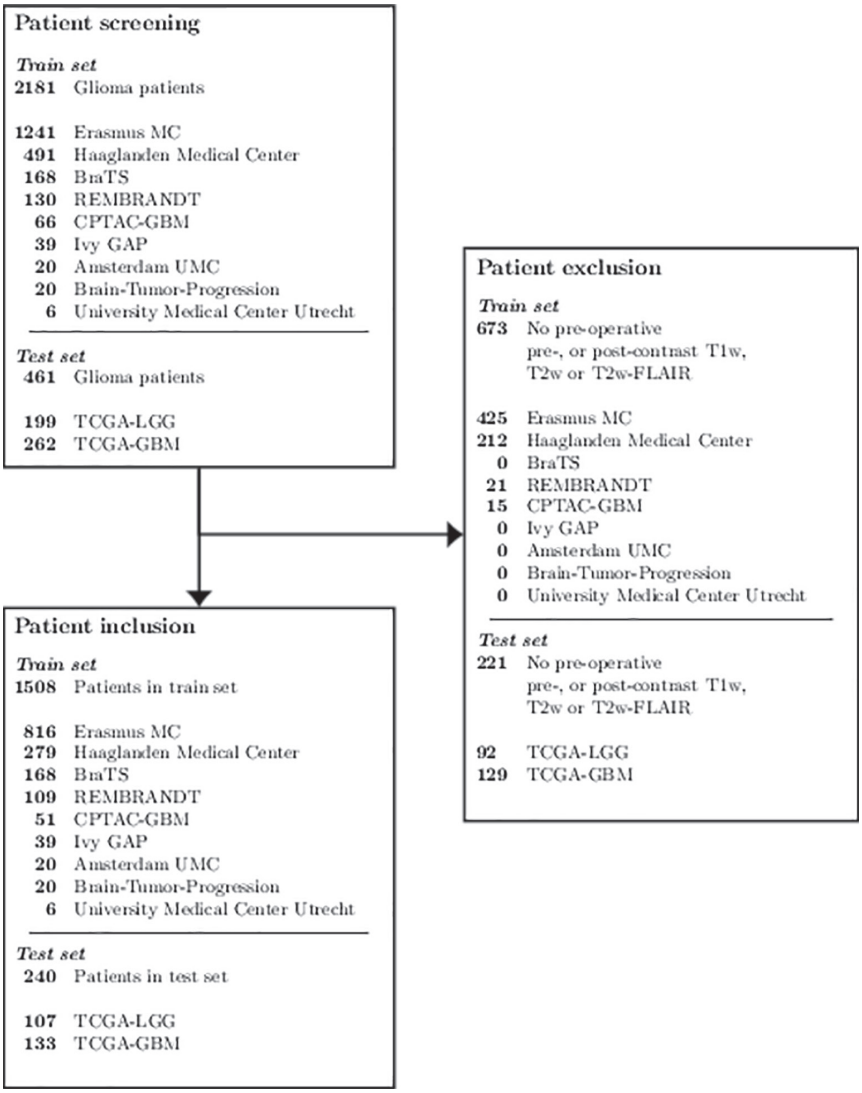


Figure 2. Inclusion flowchart of the train set and test set.

Table 1. Patient characteristics for the train set and test set.

	Train set		Test set	
	N	%	N	%
Patients	1508		240	
IDH status				
Mutated	226	15.0	88	36.7
Wildtype	440	29.2	129	53.7
Unknown	842	55.8	23	9.6
1p/19q co-deletion status				
Co-deleted	103	6.8	26	10.8
Intact	337	22.4	207	86.3
Unknown	1068	70.8	7	2.9
Grade				
II	230	15.3	47	19.6
III	114	7.6	59	24.6
IV	830	55.0	132	55.0
Unknown	334	22.1	2	0.8
WHO 2016 categorization				
Oligodendroglioma	96	6.4	26	10.8
Astrocytoma, IDH wildtype	31	2.1	22	9.2
Astrocytoma, IDH mutated	98	6.4	57	23.7
GBM, IDH wildtype	331	21.9	106	44.2
GBM, IDH mutated	16	1.1	5	2.1
Unknown	936	62.1	24	10.0
Segmentation				
Manual	716	47.5	240	100
Automatic	792	52.5	0	0

IDH: isocitrate dehydrogenase, WHO: World Health Organization, GBM: Glioblastoma.

Algorithm performance

We used 15% of the train set as a validation set and selected the model parameters that achieved the best performance on this validation set, where the model achieved an area under receiver operating characteristic curve (AUC) of 0.88 for the IDH mutation status prediction, an AUC of 0.76 for the 1p/19q co-deletion prediction, an AUC of 0.75 for the grade prediction, and a mean segmentation DICE score of 0.81. The selected model parameters are shown in Appendix F. We then trained a model using these parameters and the full train set, and evaluated it on the independent test set.

For the genetic and histological feature predictions, we achieved an AUC of 0.90 for the IDH mutation status prediction, an AUC of 0.85 for the 1p/19q co-deletion prediction, and an AUC of 0.81 for the grade prediction, in the test set. The full results are shown in

Table 2, with the corresponding receiver operating characteristic (ROC)-curves in Figure 3.

Table 2 also shows the results in (clinically relevant) subgroups of patients. This shows that we achieved an IDH-AUC of 0.81 in low grade glioma (LGG) (grade II/III), an IDH-AUC of 0.64 in high grade glioma (HGG) (grade IV), and a 1p/19q-AUC of 0.73 in LGG. When only predicting LGG vs. HGG instead of predicting the individual grades, we achieved an AUC of 0.91. In Appendix A we provide confusion matrices for the IDH, 1p/19q, and grade predictions, as well as a confusion matrix for the final WHO 2016 subtype, which shows that only one patient was predicted as a non-existing WHO 2016 subtype. In Appendix C we provide the individual predictions and ground truth labels for all patients in the test set to allow for the calculation of additional metrics.

For the automatic segmentation, we achieved a mean DICE score of 0.84, a mean Hausdorff distance of 18.9 mm, and a mean volumetric similarity coefficient of 0.90. Figure 4 shows boxplots of the DICE scores, Hausdorff distances, and volumetric similarity coefficients for the different patients in the test set. In Appendix B we show five patients that were randomly selected from both the TCGA-LGG and TCGA-GBM data collections, to demonstrate the automatic segmentations made by our method.

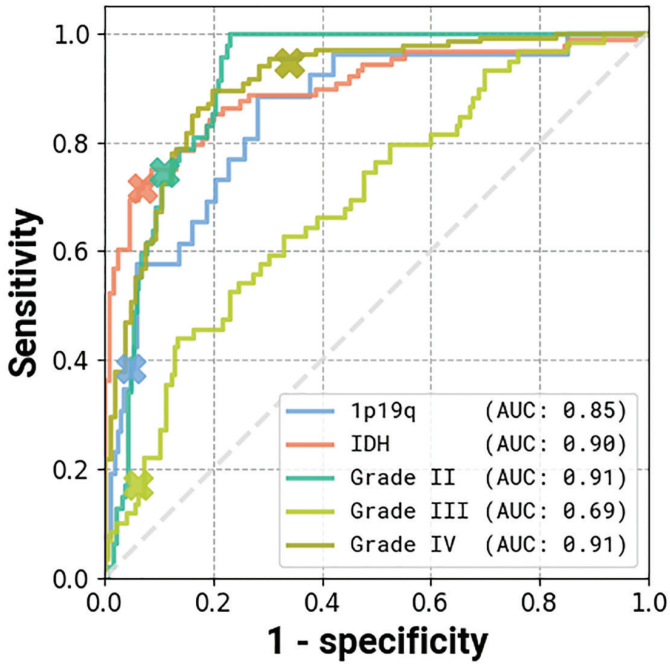


Figure 3. ROC curves of the genetic and histological features, evaluated on the test set. The crosses indicate the location of the decision threshold for the reported accuracy, sensitivity, and specificity.

Table 2. Evaluation results of the model on the test set.

Patient group	Task	AUC	Accuracy	Sensitivity	Specificity
ALL	IDH	0.90	0.84	0.72	0.93
	1p/19q	0.85	0.89	0.39	0.95
	Grade (II/III/IV)	0.81	0.71	N/A	N/A
	Grade II	0.91	0.86	0.75	0.89
	Grade III	0.69	0.75	0.17	0.89
	Grade IV	0.91	0.82	0.95	0.66
	LGG vs HGG	0.91	0.84	0.72	0.93
LGG	IDH	0.81	0.74	0.73	0.77
	1p/19q	0.81	0.74	0.73	0.77
HGG	IDH	0.64	0.94	0.40	0.96

AUC: area under receiver operating operating characteristic curce, IDH: isocitrate dehydrogenase, LGG: low grade glioma, HGG: high grade glioma..

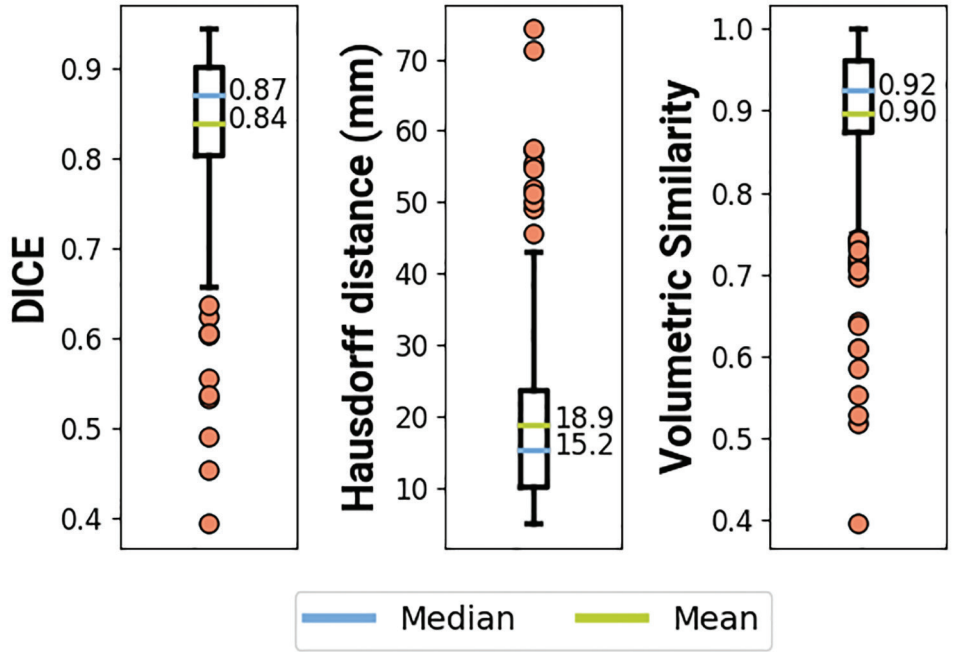
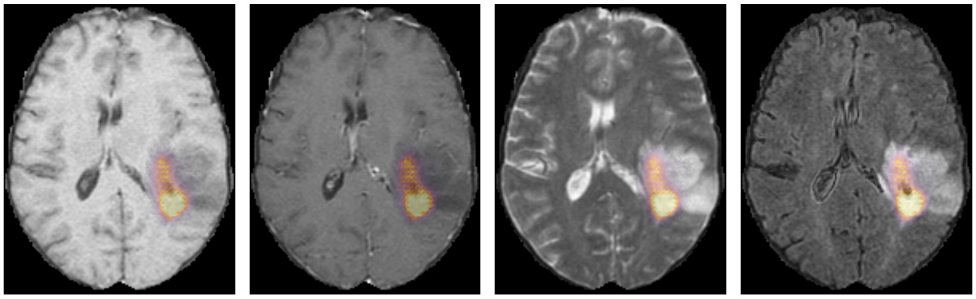


Figure 4. DICE scores, Hausdor_ distances, and volumetric similarity coefficients for all patients in the test set. The DICE score is a measure of the overlap between the ground truth and predicted segmentation (where 1 indicates perfect overlap). The Hausdor distance is a measure of the agreement between the boundaries of the ground truth and predicted segmentation (lower is better). The volumetric similarity coefficient is a measure of the agreement in volume (where 1 indicates perfect agreement).

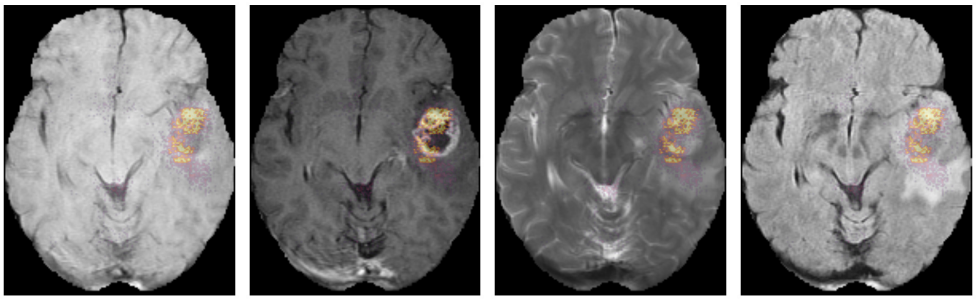
Model interpretability

To provide insight into the behavior of our model we created saliency maps, which show which parts of the scans contributed the most to the prediction. These saliency maps are shown in Figure 5 for two example patients from the test set. It can be seen that for the LGG the network focused on a bright rim in the T2w-FLAIR scan, whereas for the HGG it focused on the enhancement in the post-contrast T1w scan. To aid further interpretation, we provide visualizations of selected filter outputs in the network in Appendix D, which also show that the network focuses on the tumor, and these filters seem to recognize specific imaging features such as the contrast enhancement and T2w-FLAIR brightness.



Pre-contrast T1w | Post-contrast T1w | T2w | T2w-FLAIR

(a) Saliency maps of a low grade glioma patient (TCGA-DU-6400). This is an IDH mutated, 1p/19q co-deleted, grade II tumor. The network focuses on a rim of brightness in the T2w-FLAIR scan.



Pre-contrast T1w | Post-contrast T1w | T2w | T2w-FLAIR

(b) Saliency maps of a high grade glioma patient (TCGA-06-0238). This is an IDH wildtype, grade IV tumor. The network focuses on enhancing spots around the necrosis on the post-contrast T1w scan.

Figure 5. Saliency maps of two patients from the test set, showing areas that are relevant for the prediction.

Model robustness

By not excluding scans from our train set based on radiological characteristics, we were able to make our model robust to low scan quality, as can be seen in an example from the test set in Figure 6. Even though this example scan contained imaging artifacts, our method was able to properly segment the tumor (DICE score of 0.87), and correctly predict the tumor as an IDH wildtype, grade IV tumor.

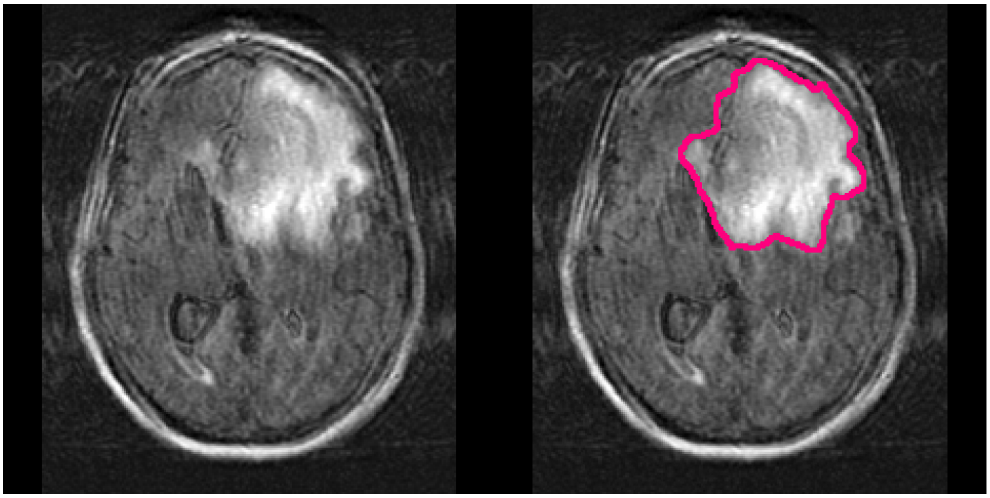
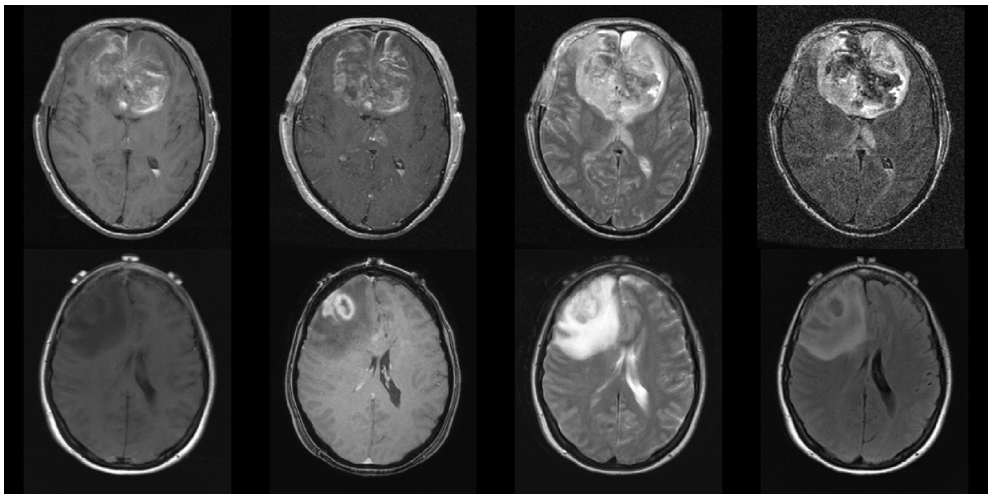


Figure 6. Example of a T2w-FLAIR scan containing imaging artifacts, and the automatic segmentation (red overlay) made by our method. It was correctly predicted as an IDH wildtype, grade IV glioma. This is patient TCGA-06-5408 from the TCGA-GBM collection.

Finally, we considered two examples of scans that were incorrectly predicted by our method, see Figure 7. These two examples were chosen because our network assigned high prediction scores to the wrong classes for these cases. Figure 7a shows an example of a grade II, IDH mutated, 1p/19q co-deleted glioma that was predicted as grade IV, IDH wildtype by our method. Our method's prediction was most likely caused by the hyperintensities in the postcontrast T1w scan being interpreted as contrast enhancement. Since these hyperintensities are also present in the pre-contrast T1w scan they are most likely calcifications, and the radiological appearance of this tumor is indicative of an oligodendroglioma. Figure 7b shows an example of a grade IV, IDH wildtype glioma that was predicted as a grade III, IDH mutated glioma by our method.



Pre-contrast T1w Post-contrast T1w T2w T2w-FLAIR

(a) TCGA-DU-6410 from the TCGA-LGG collection. The ground truth histopathological analysis indicated this glioma was grade II, IDH mutated, 1p/19q co-deleted, but our method predicted it as a grade IV, IDH wildtype. (b) TCGA-76-7664 from the TCGA-HGG collection. Histopathologically this glioma was grade IV, IDH wildtype, but our method predicted it as grade III, IDH mutated.

Figure 7. Examples of scans that were incorrectly predicted by our method.

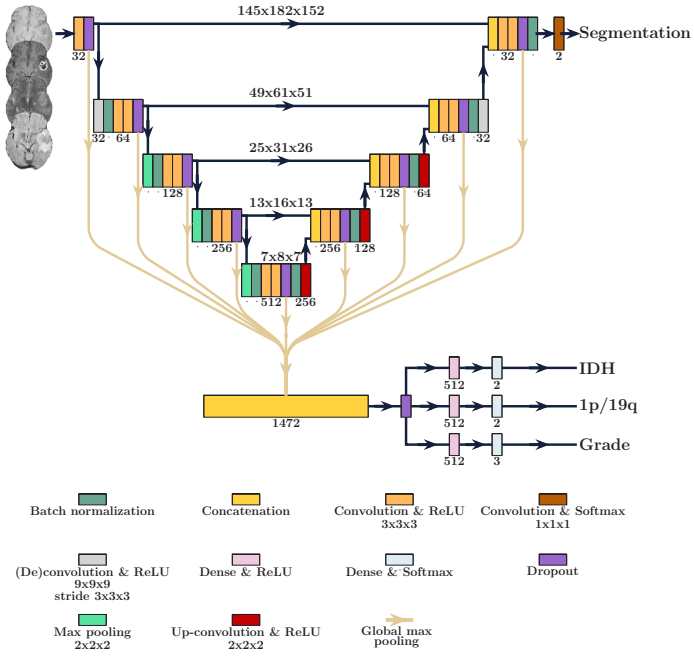


Figure 8. Overview of the PrognosAIs Structure-Net (PS-Net) architecture used for our model. The numbers below the di_erent layers indicate the number of filters, dense units or features at that layer. We have also indicated the feature map size at the different depths of the network.

Discussion

We have developed a method that can predict the IDH mutation status, 1p/19q co-deletion status, and grade of glioma, while simultaneously providing the tumor segmentation, based on pre-operative MRI scans. For the genetic and histological feature predictions, we achieved an AUC of 0.90 for the IDH mutation status prediction, an AUC of 0.85 for the 1p/19q co-deletion prediction, and an AUC of 0.81 for the grade prediction, in the test set.

In an independent test set, which contained data from 13 different institutes, we demonstrated that our method predicts these features with good overall performance; we achieved an AUC of 0.90 for the IDH mutation status prediction, an AUC of 0.85 for the 1p/19q co-deletion prediction, and an AUC of 0.81 for the grade prediction, and a mean whole tumor DICE score of 0.84. This performance on unseen data that was only used during the final evaluation of the algorithm, and that was purposefully not used to guide any decisions regarding the method design, shows the true generalizability of our method. Using the latest GPU capabilities we were able to train a large model, which uses the full 3D scan as input. Furthermore, by using the largest, most diverse patient cohort to date we were able to make our method robust to the heterogeneity that is naturally present in clinical imaging data, such that it generalizes for broad application in clinical practice.

By using a multi-task network, our method could learn the context between different features. For example, IDH wildtype and 1p/19q co-deletion are mutually exclusive [21]. If two separate methods had been used, one to predict the IDH status and one to predict the 1p/19q co-deletion status, an IDH wildtype glioma might be predicted to be 1p/19q co-deleted, which does not stroke with the clinical reality. Since our method learns both of these genetic features simultaneously, it correctly learned not to predict 1p/19q co-deletion in tumors that were IDH wildtype; there was only one patient in which our algorithm predicted a tumor to be both IDH wildtype and 1p/19q co-deleted. Furthermore, by predicting the genetic and histological features individually, instead of only predicting the WHO 2016 category, it is possible to adopt updated guidelines such as cIMPACT-NOW, future-proofing our method [22].

Some previous studies also used multi-task networks to predict the genetic and histological features of glioma [23, 24, 25]. Tang et al. [23] used a multi-task network that predicts multiple genetic features, as well as the overall survival of glioblastoma. Since their method only works for glioblastoma patients, the tumor grade must be known in advance, complicating the use of their method in the pre-operative setting when tumor grade is not yet known. Furthermore, their method requires a tumor segmentation prior to application of their method, which is a time-consuming, expert task. In a study by Xue et al. [24], a multitask network was used, with a structure similar to the one proposed in this paper, to segment the tumor and predict the grade (LGG or HGG) and IDH mutation status. However, they do not predict the 1p/19q co-deletion status needed for the WHO

2016 categorization. Lastly, Decuyper et al. [25] used a multi-task network that predicts the IDH mutation and 1p/19q co-deletion 13 status, and the tumor grade (LGG or HGG). Their method requires a tumor segmentation as input, which they obtain from a U-Net that is applied earlier in their pipeline; thus, their method requires two networks instead of the single network we use in our method. These differences aside, the most important limitation of each of these studies is the lack of an independent test set for evaluating their results. It is now considered essential that an independent test set is used, to prevent an overly optimistic estimate of a method's performance [15, 26, 27, 28]. Thus, our study improves on this previous work by providing a single network that combines the different tasks, being trained on a more extensive and diverse dataset, not requiring a tumor segmentation as an input, providing all information needed for the WHO 2016 categorization, and, crucially, by being evaluated in an independent test set.

An important genetic feature that is not predicted by our method is the O6 -methylguanine-methyltransferase (MGMT) methylation status. Although the MGMT methylation status is not part of the WHO 2016 categorization, it is part of clinical management guidelines and is an important prognostic marker in glioblastoma [4]. In the initial stages of this study, we attempted to predict the MGMT methylation status; however, the performance of this prediction was poor. Furthermore, the methylation cutoff level, which is used to determine whether a tumor is MGMT methylated, shows a wide variety between institutes, leading to inconsistent results [29]. We therefore opted not to include the MGMT prediction at all, rather than to provide a poor prediction of an unsharply defined parameter. Although some methods attempted to predict the MGMT status, with varying degrees of success, there is still an ongoing discussion on the validity of MR imaging features of the MGMT status [23, 30, 31, 32, 33].

Our method shows good overall performance, but there are noticeable performance differences between tumor categories. For example, when our method predicts a tumor as an IDH wildtype glioblastoma, it is correct almost all of the time. On the other hand, it has some difficulty differentiating IDH mutated, 1p/19q co-deleted low-grade glioma from other low-grade glioma. The sensitivity for the prediction of grade III glioma was low, which might be caused by the lack of a central pathology review. Because of this, there were differences in molecular testing and histological analysis, and it is known that distinguishing between grade II and grade III has a poor observer reliability [34]. Although our method can be relevant for certain subgroups, our method's performance still needs to be improved to ensure relevancy for the full patient population.

In future work, we aim to increase the performance of our method by including perfusion-weighted imaging (PWI) and diffusion-weighted imaging (DWI) since there has been an increasing amount of evidence that these physiological imaging modalities contain additional information that correlates with the tumor's genetic status and aggressiveness [35, 36]. They were not included in this study since PWI and, to a lesser extent, DWI are not as ingrained in the clinical imaging routine as the structural scans used in this

work [19, 20]. Thus, including these modalities would limit our method's clinical applicability and substantially reduce the number of patients in the train and test set. However, 14 PWI and DWI are increasingly becoming more commonplace, which will allow including these in future research and which might improve performance.

In conclusion, we have developed a non-invasive method that can predict the IDH mutation status, 1p/19q co-deletion status, and grade of glioma, while at the same time segmenting the tumor, based on pre-operative MRI scans with high overall performance. Although the performance of our method might need to be improved before it will find widespread clinical acceptance, we believe that this research is an important step forward in the field of radiomics. Predicting multiple clinical features simultaneously steps away from the conventional single-task methods and is more in line with the clinical practice where multiple clinical features are considered simultaneously and may even be related. Furthermore, by not limiting the patient population used to develop our method to a selection based on clinical or radiological characteristics, we alleviate the need for a priori (expert) knowledge, which may not always be available. Although steps still have to be taken before radiomics will find its way into the clinic, especially in terms of performance, our work provides a crucial step forward by resolving some of the hurdles of clinical implementation now, and paving the way for a full transition in the future.

REFERENCES

1. OFFICE FOR NATIONAL STATISTICS. CANCER SURVIVAL IN ENGLAND: Adult, Stage at Diagnosis and Childhood-Patients Followed Up... to 2018. DANDY BOOKSELLERS Limited, 2019.
2. HJ Dubbink, PN Atmodimedjo, JM Kros, PJ French, M Sanson, A Idbaih, et al. Molecular classification of anaplastic oligodendroglioma using next-generation sequencing: a report of the prospective randomized EORTC brain tumor group 26951 phase III trial. NeuroOncology, 18(3):388–400, March 2016. ISSN 1522-8517. URL https: //doi.org/10.1093/neuonc/nov182.
3. JE Eckel-Passow, DH Lachance, AM Molinaro, KM Walsh, PA Decker, H Sicotte, et al. Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. New England Journal of Medicine, 372(26):2499–2508, June 2015. ISSN 0028- 4793. URL https://doi.org/10.1056/NEJMoa1407279.
4. DN Louis, A Perry, G Reifenberger, A von Deimling, D Figarella-Branger, WK Cavenee, et al. The 2016 world health organization classification of tumors of the central nervous system: a summary. Acta Neuropathologica, 131(6):803–820, June 2016. ISSN 0001-6322. URL https://doi.org/10.1007/s00401-016-1545-1.
5. C Chen, P Hsu, T Erich Wu, S Lee, C Chang, K Wei, et al. Stereotactic brain biopsy: Single center retrospective analysis of complications. Clinical Neurology and Neurosurgery, 111(10):835–839, December 2009. ISSN 0303-8467. URL https://doi.org/10.1016/ j.clineuro.2009.08.013.
6. RJ Jackson, GN Fuller, D Abi-Said, FF Lang, ZL Gokaslan, WM Shi, et al. Limitations of stereotactic biopsy in the initial management of gliomas. Neuro-Oncology, 3(3):193–200, July 2001. ISSN 1522-8517. URL https: //doi.org/10.1093/neuonc/3.3.193.
7. M Zhou, J Scott, B Chaudhury, L Hall, D Goldgof, KW Yeom, et al. Radiomics in brain tumor: Image assessment, quantitative feature descriptors, and machine-learning approaches. American Journal of Neuroradiology, 39(2):208–216, February 2018. ISSN 0195-6108. URL https: //doi.org/10.3174/ajnr.A5391.
8. WL Bi, A Hosny, MB Schabath, ML Giger, NJ Birkbak, A Mehrtash, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. CA: A Cancer Journal for Clinicians, 69(2):caac.21552, February 2019. ISSN 0007-9235. URL https://doi.org/10.3322/caac.21552.
9. A Chaddad, MJ Kucharczyk, P Daniel, S Sabri, BJ Jean-Claude, T Niazi, et al. Radiomics in glioblastoma: Current status and challenges facing clinical implementation. Frontiers in Oncology, 9:374–374, May 2019. ISSN 2234- 943X. URL https://doi.org/10.3389/fonc.2019.00374.
10. M Smits. Imaging of oligodendroglioma. The British Journal of Radiology, 89(1060):20150857, April 2016. ISSN 0007-1285. URL https: //doi.org/10.1259/bjr.20150857.
11. RL Delfanti, DE Piccioni, J Handwerker, N Bahrami, AP Krishnan, R Karunamuni, et al. Imaging correlates for the 2016 update on WHO classification of grade II/III gliomas: implications for IDH, 1p/19q and ATRX status. Journal of Neuro-Oncology, 135(3):601–609, December 2017. ISSN 0167-594X. URL https://doi.org/10.1007/s11060-017-2613-7.

12. S Gore, T Chougule, J Jagtap, J Saini, and M Ingalhalikar. A review of radiomics and deep predictive modeling in glioma characterization. Academic Radiology, July 2020. ISSN 1076-6332. URL https://doi.org/10.1016/j.acra.2020.06.016.
13. HJWL Aerts, ER Velazquez, RTH Leijenaar, C Parmar, P Grossmann, S Carvalho, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications, 5(1):4006, September 2014. ISSN 2041-1723. URL https://doi.org/10.1038/ncomms5006.
14. S Chihati and D Gaceb. A review of recent progress in deep learning-based methods for MRI brain tumor segmentation. In 2020 11th International Conference on Information and Communication Systems (ICICS), pages 149–154. Institute of Electrical and Elect alronics Engineers (IEEE), April 2020. ISBN 9781728162270. URL https://doi.org/ 10.1109/icics49469.2020.239550.
15. RJ Gillies, PE Kinahan, and H Hricak. Radiomics: Images are more than pictures, they are data. Radiology, 278(2):563–577, February 2016. ISSN 0033-8419. URL https://doi.org/10.1148/radiol. 2015151169.
16. JH Thrall, X Li, Q Li, C Cruz, S Do, K Dreyer, et al. Artificial intelligence and machine learning in radiology: Opportunities, challenges, pitfalls, and criteria for success. Journal 39 of the American College of Radiology, 15(3, Part B):504–508, March 2018. ISSN 1546-1440. URL https://doi.org/10.1016/j.jacr.2017.12.026.
17. A Hosny, C Parmar, J Quackenbush, LH Schwartz, and HJWL Aerts. Artificial intelligence in radiology. Nature Reviews Cancer, 18(8):500–510, August 2018. ISSN 1474-175X. URL https://doi.org/10.1038/s41568-018-0016-5.
18. O Kopuklu, N Kose, A Gunduz, and G Rigoll. Resource efficient 3D convolutional neural networks. In 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Institute of Electrical and Electronics Engineers (IEEE), October 2019. ISBN 9781728150239. URL https://doi.org/10.1109/iccvw.2019.00240.
19. SC Thust, S Heiland, A Falini, HR Jager, AD Waldman, PC Sundgren, et al. Glioma imaging in europe: A survey of 220 centres and recommendations for best clinical practice. European Radiology, 28(8):3306–3317, August 2018. ISSN 0938-7994. URL https://doi.org/ 10.1007/s00330-018-5314-5.
20. CF Freyschlag, SM Krieg, J Kerschbaumer, D Pinggera, MT Forster, D Cordier, et al. Imaging practice in low-grade gliomas among european specialized centers and proposal for a minimum core of imaging. Journal of Neuro-Oncology, 139(3):699–711, September 2018. ISSN 0167- 594X. URL https://doi.org/10.1007/s11060-018-2916-3.
21. M Labussiere, A Idbaih, X-W Wang, Y Marie, B Boisselier, C Falet, S Paris, et al. All the 1p19q codeleted gliomas are mutated on IDH1 or IDH2. Neurology, 74(23):1886– 1890, June 2010. ISSN 0028-3878. URL https://doi.org/10.1212/WNL. 0b013e3181e1cf3a.
22. DN Louis, P Wesseling, K Aldape, DJ Brat, D Capper, IA Cree, et al. cIMPACT-NOW update 6: new entity and diagnostic principle recommendations of the cIMPACT-Utrecht meeting on future

CNS tumor classification and grading. Brain Pathology, 30(4):844–856, July 2020. ISSN 1015-6305. URL https://doi.org/10.1111/bpa.12832.

23. Z Tang, Y Xu, Z Jiao, J Lu, L Jin, A Aibaidula, et al. Lecture Notes in Computer Science, pages 415–422. Springer Science and Business Media LLC, 2019. ISBN 9783030322380. URL https://doi.org/10.1007/ 978-3-030-32239-7_46.
24. Z Xue, B Xin, D Wang, and X Wang. Radiomicsenhanced multi-task neural network for non-invasive glioma subtyping and segmentation. In Hassan Mohy-ud Din and Saima Rathore, editors, Radiomics and Radiogenomics in Neuro-oncology, pages 81–90. Springer Science and Business Media LLC, 2020. ISBN 9783030401238. URL https://doi.org/10.1007/978-3-030-40124-5_9.
25. M Decuyper, S Bonte, K Deblaere, and R Van Holen. Automated MRI based pipeline for glioma segmentation and prediction of grade, IDH mutation and 1p19q co-deletion, 2020. Preprint at https: //arxiv.org/abs/2005.11965.
26. S Rizzo, F Botta, S Raimondi, D Origgi, C Fanciullo, AG Morganti, et al. Radiomics: the facts and the challenges of image analysis. European Radiology Experimental, 2(1):36, December 2018. ISSN 2509-9280. URL https://doi.org/10.1186/s41747-018-0068-z.
27. P Lohmann, N Galldiks, M Kocher, A Heinzel, CP Filss, C Stegmayr, FM Mottaghy, et al. Radiomics in neuro-oncology: Basics, workflow, and applications. Methods, June 2020. ISSN 1046-2023. URL https://doi.org/10.1016/j.ymeth.2020.06.003.
28. SSF Yip and HJWL Aerts. Applications and limitations of radiomics. Physics in Medicine and Biology, 61(13):R150–R166, July 2016. ISSN 0031-9155. URL https://doi.org/10.1088/0031-9155/61/ 13/r150.
29. A Malmstrom, MI Lysiak, BW Kristensen, E Hovey, R Henriksson, and P Soderkvist. Do we really know who has an MGMT methylated glioma? Results of an international survey 41 regarding use of MGMT analyses for glioma. Neuro-Oncology Practice, 7 (1):68–76, 09 2019. ISSN 2054-2577. URL https://doi.org/10.1093/ nop/npz039.
30. T Sasaki, M Kinoshita, K Fujita, J Fukai, N Hayashi, Y Uematsu, et al. Radiomics and MGMT promoter methylation for prognostication of newly diagnosed glioblastoma. Scientific Reports, 9(1):14435, December 2019. ISSN 2045-2322. URL https://doi.org/10.1038/s41598-019-50849-y.
31. A Gupta, A Prager, RJ Young, W Shi, AMP Omuro, and JJ Graber. Diffusion-weighted MR imaging and MGMT methylation status in glioblastoma: A reappraisal of the role of preoperative quantitative ADC measurements. American Journal of Neuroradiology, 34(1):E10–E11, January 2013. ISSN 0195-6108. URL https://doi.org/10.3174/ajnr.A3467.
32. JA Carrillo, A Lai, PL Nghiemphu, HJ Kim, HS Phillips, S Kharbanda, P Moftakhar, S Lalaezari, W Yong, et al. Relationship between tumor enhancement, edema, IDH1 mutational status, MGMT promoter methylation, and survival in glioblastoma. American Journal of Neuroradiology, 33(7):1349–1355, August 2012. ISSN 0195-6108. URL https://doi.org/10.3174/ajnr.A2950.
33. VE Mikkelsen, HY Dai, AL Stensjøen, EM Berntsen, Ø Salvesen, O Solheim, et al. MGMT promoter methylation status is not related to histological or radiological features in IDH wild-type

glioblastomas. Journal of Neuropathology & Experimental Neurology, 79(8):855–862, August 2020. ISSN 0022-3069. URL https://doi.org/10.1093/jnen/nlaa060.

34. MJ van den Bent. Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. Acta Neuropathologica, 120(3):297–304, September 2010. ISSN 1432-0533. URL https://doi.org/10.1007/s00401-010-0725-7.
35. JE Park, HS Kim, Y Jo, RE Yoo, SH Choi, et al. Radiomics prognostication model in glioblastoma using diffusion- and perfusion-weighted MRI. Scientific Reports, 10(1):4250, December 2020. ISSN 2045-2322. URL https://doi. org/10.1038/s41598-020-61178-w.
36. M Kim, SY Jung, JE Park, Y Jo, SY Park, SJ Nam, et al. Diffusionand perfusion-weighted MRI radiomics model may predict isocitrate dehydrogenase (IDH) mutation and tumor aggressiveness in diffuse lower grade 42 glioma. European Radiology, 30(4):2142–2151, April 2020. ISSN 0938-7994. URL https://doi.org/10.1007/s00330-019-06548-3.
37. M Visser, DMJ Muller, RJM van Duijn, M Smits, N Verburg, EJ Hendriks, et al. Inter-rater agreement in glioma segmentations on longitudinal MRI. NeuroImage: Clinical, 22:101727, 2019. ISSN 2213-1582. URL https://doi.org/10.1016/j.nicl.2019.101727.
38. K Clark, B Vendt, K Smith, J Freymann, J Kirby, P Koppel, et al. The cancer imaging archive (TCIA): Maintaining and operating a public information repository. Journal of Digital Imaging, 26(6):1045–1057, December 2013. ISSN 0897-1889. URL https://doi.org/10.1007/s10278-013-9622-7.
39. L Scarpace, AE Flanders, R Jain, T Mikkelsen, and DW Andrews. Data from REMBRANDT. The Cancer Imaging Archive, 2015. URL https://doi.org/10.7937/K9/TCIA.2015.588OZUZB.
40. National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). Radiology data from the clinical proteomic tumor analysis consortium glioblastoma multiforme CPTAC-GBM collection. The Cancer Imaging Archive, 2018. URL https://doi.org/10.7937/k9/tcia.2018.3rje41q1.
41. N Shah, X Feng, M Lankerovich, RB Puchalski, and B Keogh. Data from Ivy GAP. The Cancer Imaging Archive, 2016. URL https://doi.org/10.7937/K9/TCIA.2016.XLwaN6nL.
42. RB Puchalski, N Shah, J Miller et al. An anatomic transcriptional atlas of human glioblastoma. Science, 360(6389):660–663, May 2018. ISSN 0036-8075. URL https://doi.org/10.1126/science.aaf2666.
43. K Schmainda and M Prah. Data from Brain-TumorProgression. The Cancer Imaging Archive, 2018. URL https://doi.org/ 10.7937/K9/TCIA.2018.15quzvnb.
44. B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, et al. The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Transactions on Medical Imaging, 34(10):1993–2024, 2015. URL https: //doi.org/10.1109/TMI.2014.2377694.
45. S Bakas, H Akbari, A Sotiras, M Bilello, M Rozycki, JS Kirby, et al. Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. Scientific Data, 4 (1):170117, December 2017. ISSN 2052-4463. URL https://doi.org/10. 1038/sdata.2017.117.

46. S Bakas, M Reyes, A Jakab, et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge, 2018. Preprint at https://arxiv.org/abs/1811.02629.

47. N Pedano, AE Flanders, L Scarpace, T Mikkelsen, JM Eschbacher, B Hermes, et al. Radiology data from the cancer genome atlas low grade glioma [TCGA-LGG] collection. The Cancer Imaging Archive, 2016. URL https://doi.org/10.7937/K9/TCIA.2016.L4LTD3TK.

48. L Scarpace, T Mikkelsen, S Cha, S Rao, S Tekchandani, D Gutman, et al. Radiology data from the cancer genome 46 atlas glioblastoma multiforme [TCGA-GBM] collection. The Cancer Imaging Archive, 2016. URL https://doi.org/10.7937/K9/TCIA.2016. RNYFUYE9.

49. S Bakas, H Akbari, A Sotiras, M Bilello, M Rozycki, J Kirby, et al. Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection [data set] The Cancer Imaging Archive, 2017. URL https://doi.org/10.7937/K9/TCIA.2017.GJQ7R0EF.

50. S Bakas, H Akbari, A Sotiras, M Bilello, M Rozycki, J Kirby, et al. Segmentation labels and radiomic features for the preoperative scans of the TCGA-GBM collection [data set]. The Cancer Imaging Archive, 2017. URL https://doi.org/10.7937/K9/TCIA.2017. KLXWJJ1Q.

51. X Li, PS Morgan, J Ashburner, J Smith, and C Rorden. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. Journal of Neuroscience Methods, 264:47–56, May 2016. ISSN 0165-0270. URL https://doi.org/10.1016/j.jneumeth.2016.03. 001.

52. V Fonov, AC Evans, K Botteron, CR Almli, RC McKinstry, and DL Collins. Unbiased average age-appropriate atlases for pediatric studies. NeuroImage, 54(1):313–327, January 2011. ISSN 1053-8119. URL https://doi.org/10.1016/j.neuroimage.2010. 07.033.

53. VS Fonov, AC Evans, RC McKinstry, CR Almli, and DL Collins. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. NeuroImage, 47:S102, July 2009. ISSN 1053-8119. URL https://doi. org/10.1016/S1053-8119(09)70884-5. Organization for Human Brain Mapping 2009 Annual Meeting.

54 S Klein, M Staring, K Murphy, MA Viergever, and J Pluim. elastix: A toolbox for intensity-based medical image registration. IEEE Transactions on Medical Imaging, 29(1):196–205, January 2010. ISSN 0278-0062. URL https://doi.org/10.1109/TMI.2009.2035616.

55 D Shamonin. Fast parallel image registration on CPU and GPU for diagnostic classification of alzheimer's disease. Frontiers in Neuroinformatics, 7:50, 2013. ISSN 1662-5196. URL https://doi. org/10.3389/ fninf.2013.00050.

56. BC Lowekamp, DT Chen, L Ibanez, and D Blezek. The design of SimpleITK. Frontiers in Neuroinformatics, 7:45, 2013. ISSN 1662-5196. URL https://doi.org/10.3389/fninf.2013.00045. 47

57. F Isensee, M Schell, I Pflueger, G Brugnara, D Bonekamp, U Neuberger, et al. Automated brain extraction of multisequence MRI using artificial neural networks. Human Brain Mapping, 40(17):4952–4964, December 2019. ISSN 1065-9471. URL https://doi.org/10.1002/hbm. 24750.

58. O Ronneberger. Invited talk: U-net convolutional networks for biomedical image segmentation. In geb Fritzsche K Maier-Hein, geb Lehmann T Deserno, Heinz Handels, and Thomas Tolxdorff, editors, Bildverarbeitung f¨ur die Medizin 2017. Informatik aktuell, pages 3 3. Springer Science

and Business Media LLC, 2017. ISBN 9783662543443. URL https: //doi.org/10.1007/978-3-662-54345-0_3.

59. M Abadi, P Barham, J Chen, Z Chen, A Davis, J Dean, et al. TensorFlow: A system for large-scale machine learning. In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), pages 265–283. USENIX Association, November 2016. ISBN 978-1-931971-33-1. URL https://www.usenix.org/conference/osdi16/ technical-sessions/ presentation/abadi.
60. D Das, N Mellempudi, D Mudigere, D Kalamkar, S Avancha, K Banerjee, et al. Mixed precision training of convolutional neural networks using integer operations. In International Conference on Learning Representations, 2018. URL https://openreview.net/forum? id=H135uzZ0-.
61 SL. Smith, PJ Kindermans, and QV Le. Don't decay the learning rate, increase the batch size. In International Conference on Learning Representations, 2018. URL https://openreview.net/ forum? id=B1Yy1BxCZ.
62. C Woolley. NCCL: accelarated multi-GPU collective communications. URL https://images. nvidia.com/events/sc15/pdfs/NCCL-Woolley. pdf. Accessed on 2020-09-30.
63. I Loshchilov and F Hutter. Decoupled weight decay regularization, 2017. Preprint at https:// arxiv.org/abs/1711.05101. 48
64. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, et al. Scikitlearn: Machine learning in python. Journal of Machine Learning Research, 12:2825–2830, 2011. URL https:// www.jmlr.org/papers/v12/ pedregosa11a.html.
65. AA Taha and A Hanbury. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. BMC Medical Imaging, 15(1):29, August 2015. ISSN 1471-2342. URL https:// doi.org/10.1186/ s12880-015-0068-x.
66. D Smilkov, N Thorat, B Kim, F Vigas, and M Wattenberg. SmoothGrad: removing noise by adding noise, 2017. Preprint at https://arxiv.org/abs/1706.03825.
67. A Tharwat. Classification assessment methods. Applied Computing and Informatics, 2018. ISSN 2210-8327. URL https://doi.org/10.1016/j. aci.2018.08.003.
68. DJ Hand and RJ Till. A simple generalisation of the area under the ROC curve for multiple class classification problems. Machine Learning, 45(2):171–186, November 2001. ISSN 1573-0565. URL https: //doi.org/10.1023/A:1010920819831.

Supplementary materials

A. Confusion matrices

Tables 3, 4 and 5 show the confusion matrices for the IDH, 1p/19q, and grade predictions, and Table 6 shows the confusion matrix for the WHO 2016 subtypes.

Table 4 shows that the algorithm mainly has difficulty recognizing 1p/19q co-deleted tumors, which are mostly predicted as 1p/19q intact. Table 5 shows that most of the incorrectly predicted grade III tumors are predicted as grade IV tumors.

Table 3. Confusion matrix of the IDH predictions.

		Predicted	
		Wildtype	**Mutated**
Actual	**Wildtype**	120	9
	Mutated	26	63

Table 4. Confusion matrix of the 1p/19q predictions.

		Predicted	
		Intact	**Co-deleted**
Actual	**Intact**	197	10
	Co-deleted	16	10

Table 5. Confusion matrix of the grade predictions.

		Predicted		
		Grade II	**Grade III**	**Grade IV**
Actual	**Grade II**	35	6	6
	Grade III	19	10	30
	Grade IV	2	5	125

Table 6. Confusion matrix of the WHO 2016 predictions. The 'other' category indicates patients that were predicted as a non-existing WHO 2016 subtype, for example IDH wildtype, 1p/19q co-deleted tumors. Only one patient (TCGA-HT-A5RC) was predicted as a non-existing category. It was predicted as an IDH wildtype, 1p/19q co-deleted, grade IV tumor.

		Predicted					
		Oligoden-droglioma	**IDH-mutated astrocytoma**	**IDH-wildtype astrocytoma**	**IDH-mutated glioblastoma**	**IDH-wildtype glioblastoma**	**Other**
Actual	Oligoden-droglioma	10	8	1	0	7	0
	IDH-mutated astrocytoma	6	34	4	3	10	0
	IDH-wildtype astrocytoma	1	2	3	2	13	1
	IDH-mutated glioblastoma	0	1	0	0	3	0
	IDH-wildtype glioblastoma	0	3	3	1	96	0

Oligodendroglioma are IDH-mutated, 1p/19q co-deleted, grade II/III gioma.
IDH-mutated astrocytoma are IDH-mutated, 1p/19q intact, grade II/III glioma.
IDH-wildtype astrocytoma are IDH-wildtype, 1p/19q intact, grade II/III glioma.
IDH-mutated glioblastoma are IDH-mutated, grade IV glioma.
IDH-wildtype glioblastoma are IDH-wildtype, grade IV glioma.

Table 6 shows that our algorithm often incorrectly predicts IDH-wildtype astrocytoma as IDH-wildtype glioblastoma. The latest cIMPACT-NOW guidelines propose a new categorization, in which IDH-wildtype astrocytoma that show either TERT promoter methylation, or EFGR gene amplification, or chromosome 7 gain/chromsome 10 loss are classified as IDH-wildtype glioblastoma [22]. This new categorization is proposed since the survival of patients with those IDH-wildtype astrocytoma is similar to the survival of patients with IDHwildtype glioblastoma [22]. From the 13 IDH-wildtype astrocytoma that were wrongly predicted as IDH-wildtype glioblastoma, 12 would actually be categorized as IDH-wildtype glioblastoma under this new categorization. Thus, although our method wrongly predicted the WHO 2016 subtype, it might actually have picked up on imaging features related to the aggressiveness of the tumor, which might lead to a better categorization.

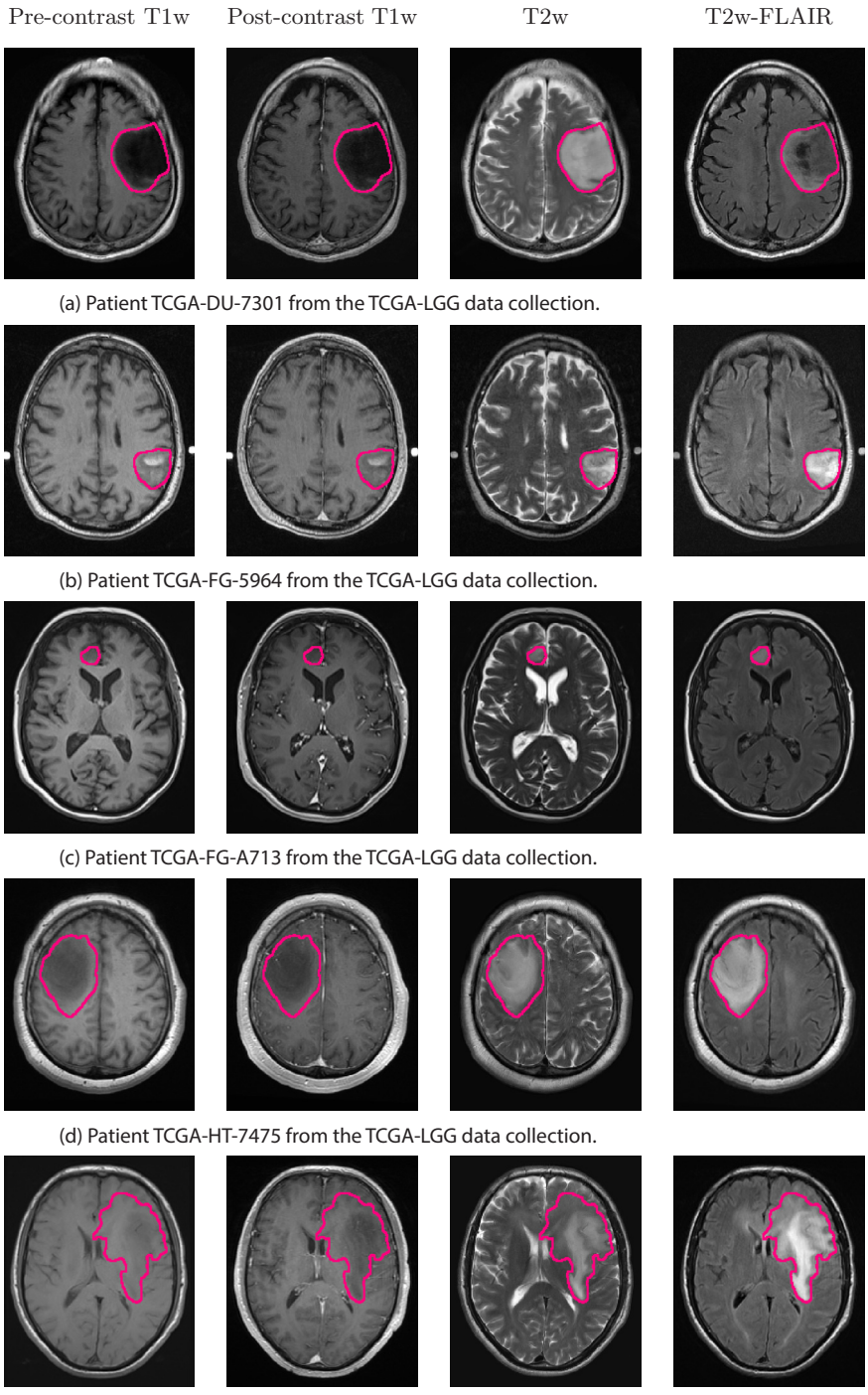
B. Segmentation examples

To demonstrate the automatic segmentations made by our method, we randomly selected five patients from both the TCGA-LGG and the TCGA-GBM dataset. The scans and segmentations of the five patients from the TCGA-LGG dataset and the TCGA-GBM dataset are shown in Figures 9 and 10, respectively. The DICE score, Hausdorff distance, and volumetric similarity coefficient for these patients are given in Table 7. The method seems to mostly focus on the hyperintensities of the T2w-FLAIR scan. Despite the

registrations issues that can be seen for the T2w scan in Figure 10d the tumor was still properly segmented, demonstrating the robustness of our method

Table 7. The DICE score, Hausdor distance (HD), and volumetric similarity coecient (VSC) for the randomly selected patients from the TCGA-LGG and TCGA-GBM data collections.

Patient	DICE	HD (mm)	VSC
TCGA-LGG			
TCGA-DU-7301	0.89	10.3	0.95
TCGA-FG-5964	0.80	5.8	0.82
TCGA-FG-A713	0.73	7.8	0.88
TCGA-HT-7475	0.87	14.9	0.90
TCGA-HT-8106	0.88	11.2	0.99
TCGA-GBM			
TCGA-02-0037	0.82	22.6	0.99
TCGA-08-0353	0.91	13.0	0.98
TCGA-12-1094	0.90	7.3	0.93
TCGA-14-3477	0.90	16.5	0.99
TCGA-19-5951	0.73	19.7	0.73



(a) Patient TCGA-DU-7301 from the TCGA-LGG data collection.

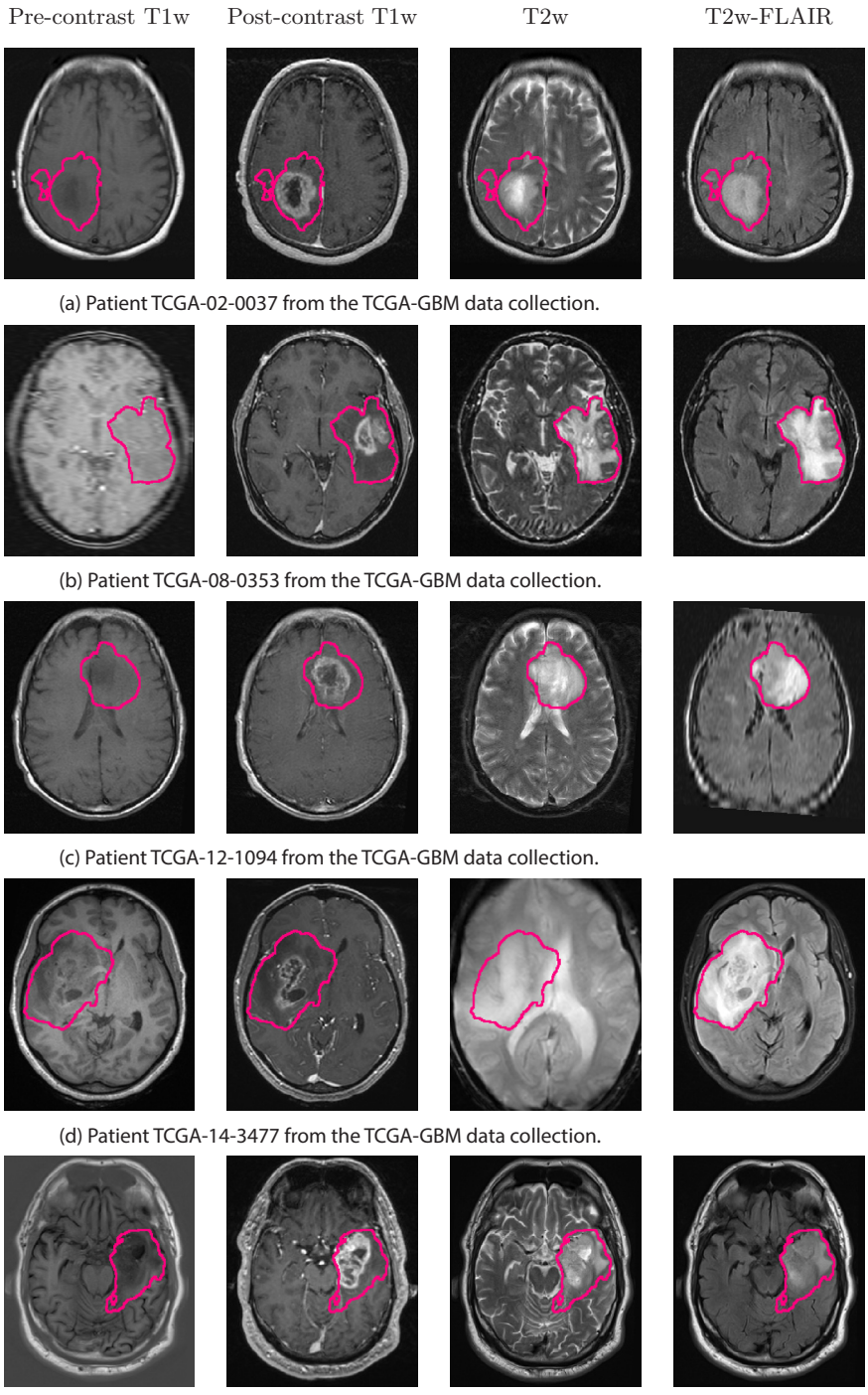
(b) Patient TCGA-FG-5964 from the TCGA-LGG data collection.

(c) Patient TCGA-FG-A713 from the TCGA-LGG data collection.

(d) Patient TCGA-HT-7475 from the TCGA-LGG data collection.

(e) Patient TCGA-HT-8106 from the TCGA-LGG data collection.

Figure 9. Examples of scans and automatic segmentations of five patients that were randomly selected from the TCCA-LGG data collection.



(a) Patient TCGA-02-0037 from the TCGA-GBM data collection.

(b) Patient TCGA-08-0353 from the TCGA-GBM data collection.

(c) Patient TCGA-12-1094 from the TCGA-GBM data collection.

(d) Patient TCGA-14-3477 from the TCGA-GBM data collection.

(e) Patient TCGA-19-5951 from the TCGA-GBM data collection.

Figure 10. Examples of scans and automatic segmentations of five patients that were randomlu selected from the TCGA-GBM data collection.

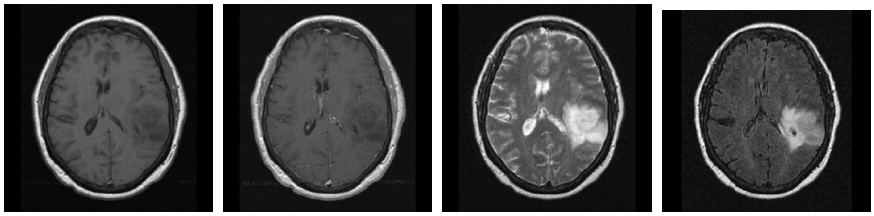
C. Prediction results in the test set

Available as online supplementary file

D. Filter output visualizations

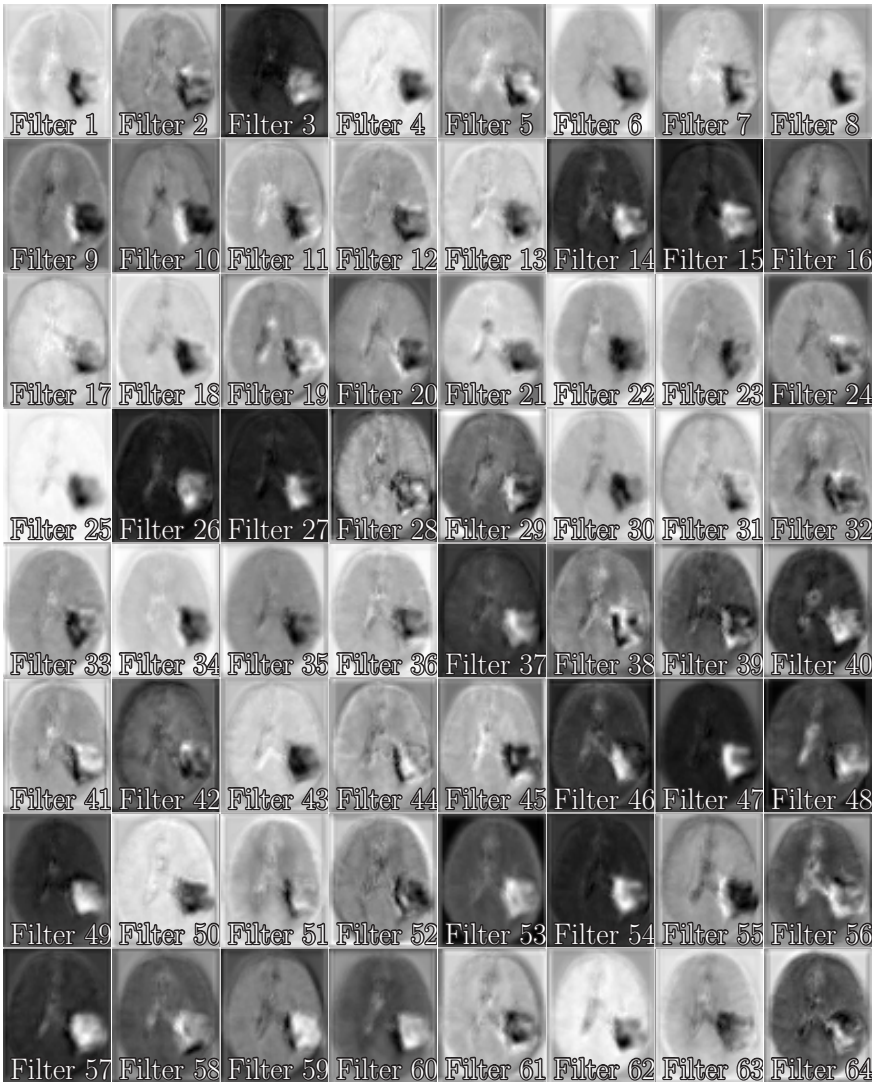
Figures 11 and 12 show the output of the convolution filters for the same LGG patient as shown in Figure 5a, and Figures 13 and 14 show the output of the convolution filters for the same HGG patient as shown in Figure 5b. Figures 11 and 13 show the outputs of the last convolution layer in the downsample path at the feature size of 49x61x51 (the fourth convolutional layer in the network). Figures 12 and 14 show the outputs of the last convolution layer in the upsample path at the feature size of 49x61x51 (the nineteenth convolutional layer in the network).

Comparing Figure 11 to Figure 12 and Figure 13 to Figure 14 we can see that the convolutional layers in the upsample path do not keep a lot of detail for the healthy part of the brain, as this region seems blurred. However, within the tumor different regions can still be distinguished. The different parts of the tumor from the scans can also be seen, such as the contrast-enhancing part and the high signal intensity on the T2w-FLAIR. For the grade IV glioma in Figure 14, some filters, such as filter 26, also seem to focus on the necrotic part of the tumor.



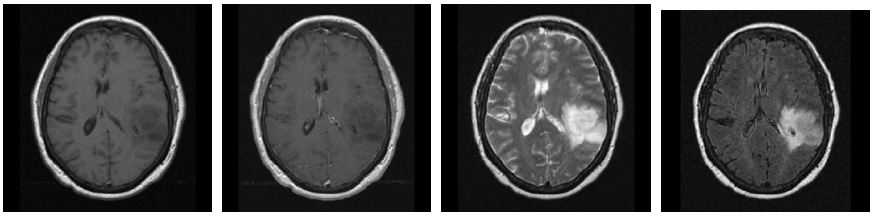
Pre-contrast T1w Post-contrast T1w T2w T2w-FLAIR

(a) Scans used to derive the convolutional layer filter output visualizations.



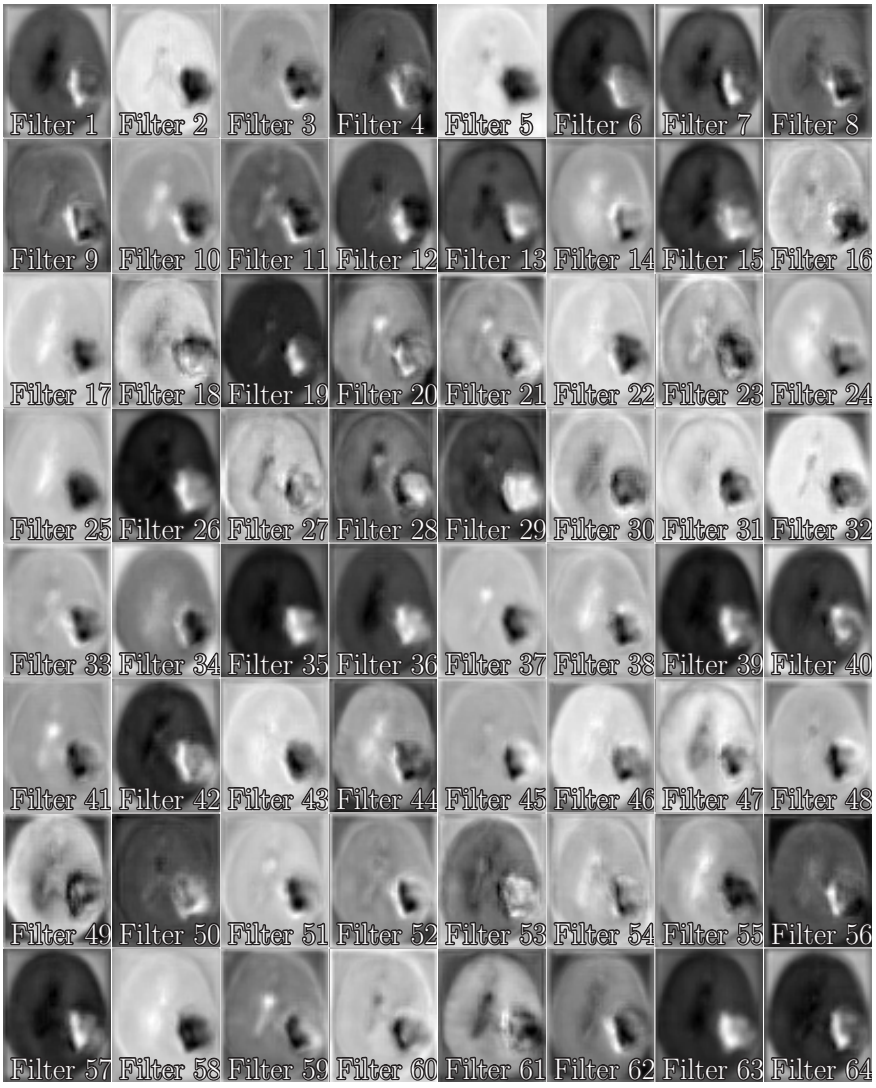
(b) Filter output visualizations.

Figure 11. Filter output visualizations of the last convolutional layer in the downsample path of the network at feature map size 49x61x51 for patient TCGA-DU-6400. This is an IDH mutated, 1p/19q co-deleted, grade II glioma.



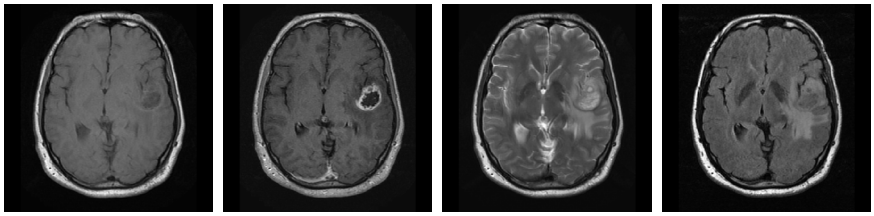
Pre-contrast T1w | Post-contrast T1w | T2w | T2w-FLAIR

(a) Scans used to derive the convolutional layer filter output visualizations.



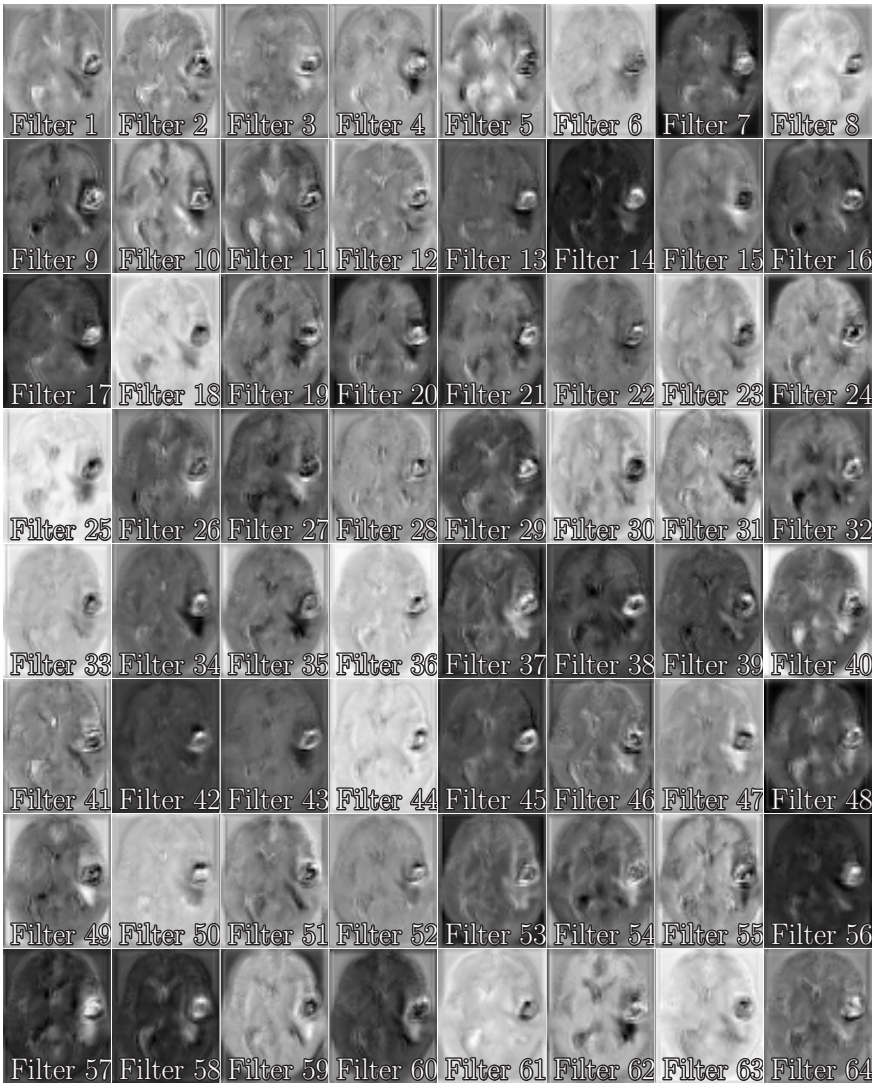
(b) Filter output visualizations.

Figure 12. Filter output visualizations of the last convolutional layer in the upsample path of the network at feature map size 49x61x51 for patient TCGA-DU-6400. This is an IDH mutated, 1p/19q co-deleted, grade II glioma.



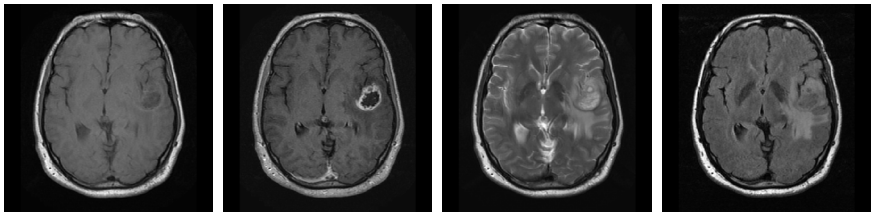
Pre-contrast T1w Post-contrast T1w T2w T2w-FLAIR

(a) Scans used to derive the convolutional layer filter output visualizations.



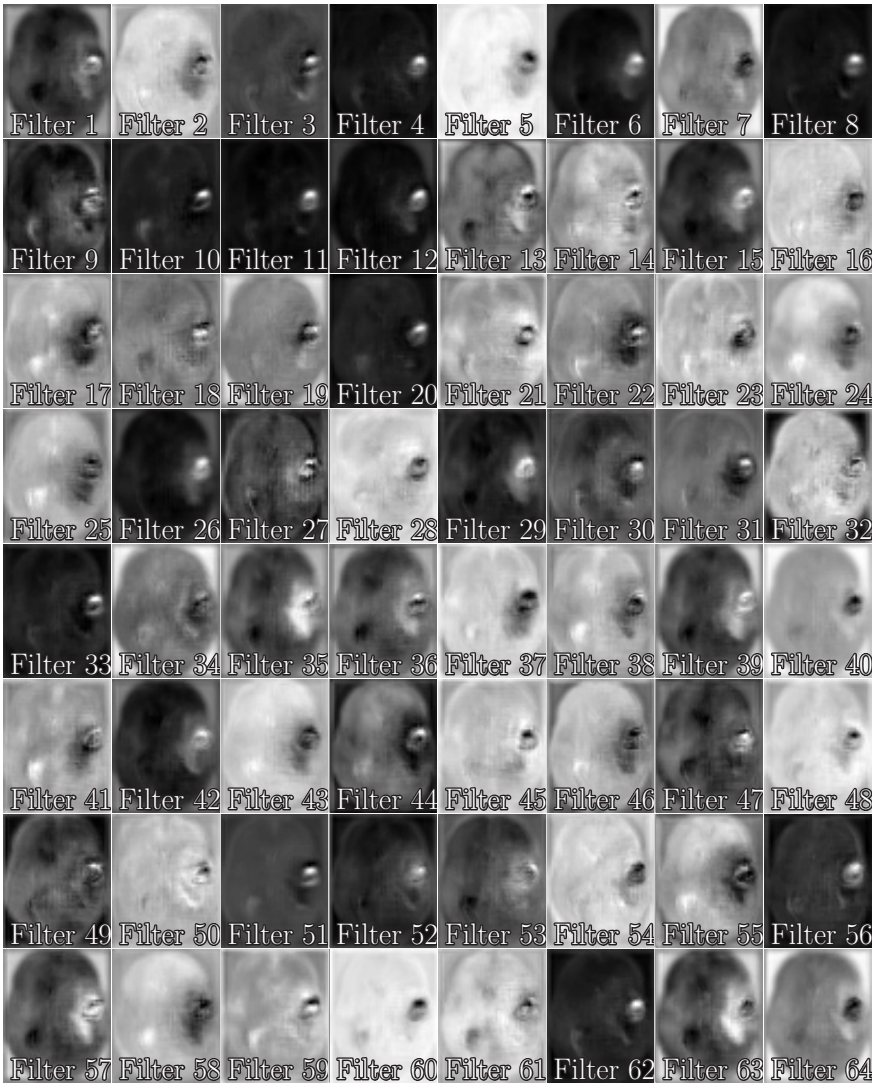
(b) Filter output visualizations.

Figure 13. Filter output visualizations of the last convolutional layer in the downsample path of the network at feature map size 49x61x51 for patient TCGA-06-0238. This is an IDH wildtype, grade IV glioma.



Pre-contrast T1w Post-contrast T1w T2w T2w-FLAIR

(a) Scans used to derive the convolutional layer filter output visualizations.



(b) Filter output visualizations.

Figure 14. Filter output visualizations of the last convolutional layer in the upsample path of the network at feature map size 49x61x51 for patient TCGA-06-0238. This is an IDH wildtype, grade IV glioma.

E. Training losses

During the training of the network we used masked categorical cross-entropy loss for the IDH, 1p/19q, and grade outputs. The normal categorical cross-entropy loss is defined as:

$$\mathcal{L}_{batch}^{CE} = -\frac{1}{N_{batch}} \sum_{j} \sum_{i \in C} y_{i,j} \log\left(\hat{y}_{i,j}\right),$$

where L CE batch is the total cross-entropy loss over a batch, yi,j is the ground truth label of sample j for class i, ˆyi,j is the prediction score for sample j for class i, C is the set of classes, and Nbatch is the number of samples in the batch. Here it is assumed that the ground truth labels are one-hot-encoded, thus yi,j is either 0 or 1 for each class. In our case, the ground truth is not known for all samples, which can be incorporated in Equation (1) by setting yi,j to 0 for all classes for a sample for which the ground truth is not known. That sample would then not contribute to the overall loss, and would not contribute to the gradient update. However, this can skew the total loss over a batch, since the loss is still averaged over the total number of samples in a batch, regardless of whether the ground truth is known, resulting in a lower loss for batches that contained more samples with unknown ground truth. Therefore, we used a masked categorical cross-entropy loss:

$$\mathcal{L}_{batch}^{CE} = -\frac{1}{N_{batch}} \sum_{j} \mu_j^{batch} \sum_{i \in C} y_{i,j} \log\left(\hat{y}_{i,j}\right),$$

$$\mu_j^{batch} = \frac{N_{batch}}{\sum_{i,j} y_{i,j}} \sum_{i} y_{i,j}$$

is the batch weight for sample j. In this way, the total batch loss is only averaged over the samples that actually have a ground truth. Since there was an imbalance between the number of ground truth samples for each class, we used class weights to compensate for this imbalance. Thus, the loss becomes:

$$\mathcal{L}_{batch}^{CE} = -\frac{1}{N_{batch}} \sum_{j} \mu_j^{batch} \sum_{i \in C} \mu_i^{class} y_{i,j} \log\left(\hat{y}_{i,j}\right),$$

$$\mu_i^{class} = \frac{N}{N_i \left|C\right|}$$

is the class weight for class i, N is the total number of samples with known ground truth, Ni is the number of samples of class i, and |C| is the number of classes. By determining the class weight in this way, we ensured that:

$$\mu_i^{class} N_i = \frac{N}{|C|} = \text{constant}.$$

Thus, each class would have the same contribution to the overall loss. These class weights were (individually) determined for the IDH output, the 1p/19q output, and the grade output. For the segmentation output we used the DICE loss:

$$\mathcal{L}_{batch}^{DICE} = \sum_j 1 - 2 \cdot \frac{\sum_k^{voxels} y_{j,k} \cdot \hat{y}_{j,k}}{\sum_k^{voxels} y_{j,k} + \hat{y}_{j,k}},$$

where yj,k is the ground truth label in voxel k of sample j, and ˆyj,k is the prediction score outputted for voxel k of sample j. The total loss that was optimized for the model was a weighted sum of the four individual losses:

$$\mathcal{L}^{total} = \sum_m \mu_m \mathcal{L}_m,$$

with

$$\mu_m = \frac{1}{X_m},$$

where Lm is the loss for output m, μm is the loss weight for loss m (either the IDH, 1p/19q, grade or segmentation loss), and Xm is the number of samples with known ground truth for output m. In this way, we could counteract the effect of certain outputs having more known labels than other outputs.

F. Parameter tuning

Table 8. Hyperparameters that were tuned, and the values that were tested. Values in bold show the selected values used in the nal model.

Tuning parameter	Tested values
Dropout rate	0.15, 0.2, **0.25**, 0.30, 0.35, 0.40
∫2-norm	0.0001, **0.00001**, 0.000001
Learning rate	0.01, 0.001, 0.0001, **0.00001**, 0.0000001
Weight decay	0.001, 0.0001, **0.00001**
Augmentation factor	1, **2**, 3
Augmentation probability	0.25, 0.30, **0.35**, 0.40, 0.45

G. Evaluation metrics

We calculated the AUC, accuracy, sensitivity, and specificity metrics for the genetic and histological features; for the definitions of these metrics see [67].

For the IDH and 1p/19q co-deletion outputs, the IDH mutated and the 1p/19q co-deleted samples were regarded as the positive class respectively. Since the grade was a multi-class problem, no single positive class could be determined. For the prediction of the individual grades, that grade was seen as the positive class and all other grades as the negative class (e.g., in the case of the grade III prediction, grade III was regarded as the positive class, and grade II and IV were regarded as the negative class). For the LGG vs. HGG prediction, LGG was considered as the positive class and HGG as the negative class. For the evaluation of these metrics for the genetic and histological features, only the subjects with known ground truth were taken into account.

The overall AUC for the grade was a multi-class AUC determined in a onevs-one approach, comparing each class against the others; in this way, this metric was insensitive to class imbalance [68]. A multi-class accuracy was used to determine the overall accuracy for the grade predictions [67].

To evaluate the performance of the automated segmentation, we evaluated the DICE score, the Hausdorff distance and the volumetric similarity coefficient. The DICE score is a measure of overlap between two segmentations, where a value of 1 indicates perfect overlap, and the Hausdorff distance is a measure of the closeness of the borders of the segmentations. The volumetric similarity coefficient is a measure of the agreement between the volumes of two segmentations, without taking account the actual location of the tumor, where a value of 1 indicates perfect agreement. See [65] for the definitions of these metrics

H. Ground truth labels of patients included from public datasets

Available as online supplementary file.

Chapter 5

Changes in language white matter tract microarchitecture are associated with cognitive deficits in patients with presumed low grade glioma

Fatih Incekara, Djaina Satoer, Evy Visch-Brink, Arnaud J.P.E. Vincent

J Neurosurg. 2018 Jun 8:1-9

Abstract

Background

The aim of this study was to determine whether cognitive functioning of presumed low-grade glioma patients is associated with white matter (WM) tract changes.

Methods

We included 77 patients with presumed low-grade glioma who underwent awake surgery between 2005 and 2013. Diffusion tensor imaging (DTI) with deterministic tractography was performed pre-operatively to identify the arcuate (AF), inferior fronto-occipital (IFOF) and uncinate fasciculus (UF) and to obtain mean fractional anisotropy (FA) and mean diffusivity per tract. All patients were assessed pre-operatively with an extensive neuropsychological protocol, including the language, memory and attention/executive function domains. Linear regression models were used per cognitive domain and per DTI metric of the three WM tracts.

Results

Significant correlations (corrected for multiple testing) were found between FA of the AF and the repetition test of the language domain ($\beta = 0.59$, $p < 0.0001$), and between FA of the IFOF and the imprinting test of the memory domain ($\beta = -0.55$, $p = 0.002$) and attention test of the attention and executive function domain ($\beta = -0.62$, $p = 0.006$).

Conclusion

In glioma patients, language deficits in repetition of speech, imprinting and attention deficits are associated with changed microarchitecture of the arcuate fasciculus and the inferior fronto-occipital fasciculus.

Introduction

Low grade gliomas (LGG) are World Health Organization (WHO) Grade I-II, slow growing, primary brain tumors that commonly show no or minimal contrast-enhancement. Patients often present with seizures, but headache and (mild) cognitive deficits are also frequently observed [31]. Due to a widespread representation of cognitive functions, it is important to make use of a neuropsychological protocol in glioma patients to detect deficits [24]. Satoer et al. showed with a neuropsychological assessment (NP) in a series of 45 glioma patients pre- and postoperative deficits in the domains of language, memory and the attention/executive functions [26]. We hypothesize that such cognitive deficits can – at least partly – be attributed to changes in white matter (WM) tract microarchitecture due to tumor effects such as edema or infiltration.

In vivo anatomical visualization of WM tracts, such as the arcuate fasciculus (AF), inferior fronto-occipital fasciculus (IFOF) and uncinate fasciculus (UF) can be obtained with the use of Diffusion Tensor Imaging (DTI) based tractography. DTI is a magnetic resonance imaging (MRI) modality in which the 3D diffusion of free water in the brain can be assessed. Since water diffuses preferentially along the direction of the white matter fibers instead of perpendicular to it, measuring the direction of free water diffusion allows for an approximation of the white matter fiber orientation. On a DTI-derived color-coded map, the white matter tracts are conventionally represented in green in the anterior-posterior direction (e.g. optical fibers), in red in the left-right direction (e.g. corpus callosum) and in blue in the ventral-dorsal direction (e.g. cortico-spinal tract). To delineate the specific white matter tracts, specific regions of interests (ROIs) are manually delineated on DTI images based on a priori anatomical knowledge. WM tracts of interest running through these ROIs are then visualized in 3D using tractography [5,32]. The AF is an association bundle that is involved in language by connecting perisylvian language areas of the frontal, temporal and parietal lobes [5]. Intraoperative stimulation of the AF elicits phonemic paraphasia [9]. The IFOF is an association bundle that connects the occipital lobe and the orbitofrontal cortex. The IFOF seems to be involved in several cognitive functions such as language (semantics and reading) [10], but also attention and visual processing abilities [5,7]. The UF is an association bundle that connects the orbitofrontal cortex with the anterior temporal lobe. The UF is suggested to be involved in memory and language. [24]

Quantitative information on WM tract microarchitecture can be derived from tractography and expressed as fractional anisotropy (FA) and mean diffusivity (MD). FA is a measure of microstructural integrity with a value between 0 and 1, with 0 meaning the diffusion is the same in all directions (isotropic diffusion, such as seen in the ventricles) and close to 1 meaning diffusion of water is seen along a specific direction (anisotropic diffusion, such as seen along white matter tracts). A reduction of FA in a white matter tract is considered to be an indication of loss of tract integrity. MD is a measure of the average water diffusion and reflects membrane density. High MD is seen in free water

(the ventricles), and low MD in the white matter. An increase of MD in the white matter is considered an indication of the increase of the extracellular space, such as seen in edema. While derived from the same imaging sequence, FA and MD reflect different aspects of white matter tract structure, and can thus be used to assess the effects of tumor infiltration and edema. [2] Importantly, perilesional WM tracts and edematous zones around glioma show a reduction in FA values which suggests increased extracellular water and fiber disorganization due to tumor infiltration[1]. Studies have shown that tumor infiltrated WM tracts such as the AF have a relationship with language impairment [3,15]. However, to the best of our knowledge there is as yet no large-scale quantitative correlation of WM tract microarchitecture with cognitive functioning in LGG patients.

The purpose of this study was to determine whether cognitive functioning of presumed LGG patients is associated with changes in language WM tract microarchitecture. We analyzed FA and MD values of the AF, IFOF, UF in relation to three cognitive domains (language, memory and attention/executive functions) assessed with an extensive NP pre-operatively in patients selected for awake surgery.

METHODS

Study participants

A total of 131 patients (> 18 years) with presumed LGG who underwent awake surgery between April 2005 and May 2013 were considered for this retrospective study. Non-to-mildly contrast enhancing tumors, i.e. presumed LGG, on pre-operative MRI scan were considered. Note that if upon histopathological examination – *after* resection – such tumors displayed high grade glioma features, patients were retained in the study, to ensure that our findings are valid for the real-life clinical situation of pre-operative assessment and selection of patients, when histopathology is not yet available. Hence the use of the term *presumed* LGG, based on clinical and radiological characteristics. All patients were native Dutch speakers. Patients with recurrent glioma (n=24), without pre-operative NP (n=3), without sufficient DTI data (n=25), or a time frame between NP and DTI longer than 6 months (n=2) were excluded, resulting in a total of 77 patients available for data analysis (Table 1). Of the 77 patients (mean age 43 years, range 20-74), 49 were male (63.6%) with a mean age of 44 years (range 21 – 74), while all 28 females (36. 4%) had a mean age of 40 years, (range 20 - 60). Although most tumors were localized in the left hemisphere, 15 patients with right-sided tumors were also included in this study, because they had right dominance for language as seen on fMRI, which was performed pre-operatively. Therefore, these patients were selected for awake surgery to prevent damage to the Broca region. The study was approved by the Medical Ethical Committee, who waived the need for written informed consent from the patients due to the retrospective nature of this study and the (emotional) burden that would result from contacting the patients or their relatives to obtain consent.

Table 1. Baseline patient characteristics.

Characteristics	N (%)	Characteristics	N (%)
Age in years		**Tumor localization**	
mean; range	43 (20 – 74)	**Right**	15 (20.0)
Gender		Frontal	3 (3.9)
male/female	49 (63.6) / 28 (36.4)	Parietal	10 (13.0)
Handedness		Temporal	2 (2.6)
Right	54 (70.1)	Occipital	0 (0)
Left	5 (6.5)	**Left**	62 (80.0)
Unknown	18 (23.4)	Frontal	12 (15.6)
Karnofsky Performance Status		Parietal	23 (29.9)
100-90	69 (90)	Temporal	26 (33.8)
80-70	8 (10)	Occipital	1 (1.3)
Tumor grade		**Pathological tumor growth**	
Low grade glioma	47 (61)	Diffuse infiltrative	37 (48.1)
High grade glioma	30 (39)	Circumscribed	40 (51.9)
Histopathological subtype			
Oligodendroglioma WHO grade II	21 (27.3)		
Astrocytoma WHO grade II	13 (16.9)		
Oligoastrocytoma WHO grade II	11 (14.3)		
Oligodendroglioma WHO grade III	16 (20.8)		
Astrocytoma WHO grade III	12 (15.6)		
Oligoastrocytoma WHO grade III	3 (3.9)		
Glioblastoma WHO grade IV	1 (1.3)		

WHO = World Health Organization.

Tumor localization and characterization

Tumor localization was determined by a neuroradiologist (8 years of experience) on pre-operative T1 and T2 weighted images. The histopathological type and grade of the tumor were determined by a neuropathologist from tissue obtained during tumor resection.

Image acquisition and preprocessing

Diffusion MR data were acquired using a single shot spin echo planar imaging (EPI) sequence at 1.5T (n = 20) and 3.0T (n = 57) (GE Healthcare, Milwaukee, USA) with an eight-

channel head coil. In general, 25 non-collinear gradient directions at b = 1000 s/mm^2 (n = 71) and 3 images at b = 0 s/mm^2 were acquired (n = 64). In six patients, all imaged at 3.0T, 31 gradient directions were acquired with 4 images at b = 0 s/mm^2. Slice thickness was 2.0-3.5 mm and in-plane resolution 1.9-3.4 mm^2. Raw diffusion MRI data were transferred to an offline workstation. All images were first visually inspected for the presence of apparent artefacts. MRI data were pre- and post-processed using ExploreDTI [17]. Motion and eddy current correction of the DWI was not performed to avoid interpolation errors. Instead, the acquired native data and the data-quality-summary report were visually inspected in ExploreDTI. The gradient components in the x/y/z axes were manually checked and adapted if needed to the standard color-convention (left-right: red, top-bottom: blue, and front-back: green). Subsequently, the diffusion tensors were estimated using nonlinear least squares. The following thresholds were used to perform tractography: FA termination threshold of 0. 2, angle threshold of 45°, step length: 0/5.

Tractography

Deterministic tractography was used to identify the AF, IFOF and UF in the affected hemisphere according to standardized procedures [30]. From each of the reconstructed tracts, average FA and MD values were obtained. The rater was blinded for NP test results of the patients at time of tractography.

Arcuate fasciculus

For the AF (Figure 1), a coronal slice was selected at the level of the primary motor cortex and the seed ROI was placed on the green (anterior-posterior) triangular shaped projections of the superior longitudinal fasciculus lateral to the blue projections of the cortical spinal tract in the posterior parietal lobe. For the target ROI, an axial slice was selected at the level of the genu of the corpus callosum. The blue (cranial-caudal) projections lateral to the sagittal stratum (green) in the posterior temporal lobe corresponding to the vertical portion of the AF were delineated [30].

Inferior fronto-occipital fasciculus

For the IFOF (Figure 2), the seed ROI was placed by delineating the green projections in the entire anterior part of the frontal lobe on the axial and/or coronal slices. The target ROI was placed by delineating the green projections in the entire posterior part of the occipital lobe up until the posterior crossing area between the genu of the corpus callosum (red) and the posterior part of the sagittal striatum (green) on the axial and/or coronal slices [30] .

Uncinate fasciculus

For the UF (Figure 3), the seed ROI was placed in the green projections in the temporal pole/ entire anterior part of the temporal lobe on the coronal slice. The target ROI was

placed by delineating the green projections in the inferior part of the frontal lobe on the coronal slice [30].

Neuropsychological assessment

All patients were assessed once with a comprehensive NP test battery after initial diagnosis (mean= 9 days; SD= 20) and before surgery (mean= 44 days; SD= 30) by a clinical linguist (blinded for tractography results). As shown in Table 2 [12,14,18,20,27,33], the performance of the patients was examined within three main cognitive domains (language, memory and attention/executive functions). This combination of NP tests has shown to be sensitive to detect cognitive deficits and has been validated in glioma patients in earlier studies [24,26,31]. All test-scores of the patients were transformed into z-values to compare the performance of patients to a normative group and to facilitate comparisons between tests.

Statistical Analysis

Statistical analyses were performed with the SPSS statistical software (SPSS version 21.0, SPSS Inc. Chicago, IL, USA). We determined whether patients' mean scores by cognitive domain and by test differed from the normative group, using either a one-sample t-test with 0 (the mean score of the normal group) as test value or the Wilcoxon Signed Rank Test. Post-hoc, mean FA/MD of the three WM tracts and three cognitive domains were compared between histopathological established LGG and HGG by using a one-way ANOVA. First the relationship between the three main cognitive domains (language, memory and attention/executive functions) of the NP and FA/MD values of three WM tracts (AF, IFOF, UF) was analyzed in linear regression models which included all three WM tracts, with age and histopathological grade as regressors of no interest. The analyses were done separately for each main cognitive domain and each DTI metric (FA, MD). Bonferroni correction was used to adjust for multiple testing on main cognitive domain level (6 tests): a p value of ≤ 0.008 was considered to be statistically significant. Then, further analyses were done on subtest level, only within the cognitive domain(s) that were significantly correlated with WM tract measure(s). Bonferroni correction was not used for these analyses: a p value of <0. 05 was considered to be significant.

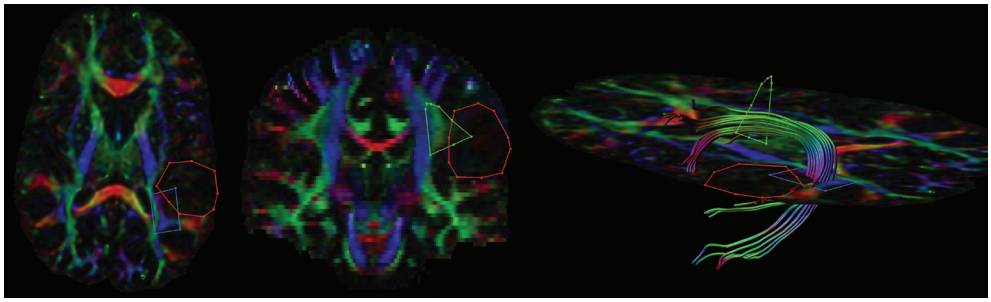


Figure 1. The first ROI is placed on the green triangular shaped projections lateral to the blue projections of the cortical spinal tract in the posterior parietal lobe. The second ROI is placed on the blue projections laterally to the sagittal stratum (green). The tumor is delineated in red. The 3D-reconstruction of the AF is shown at the most right figure.

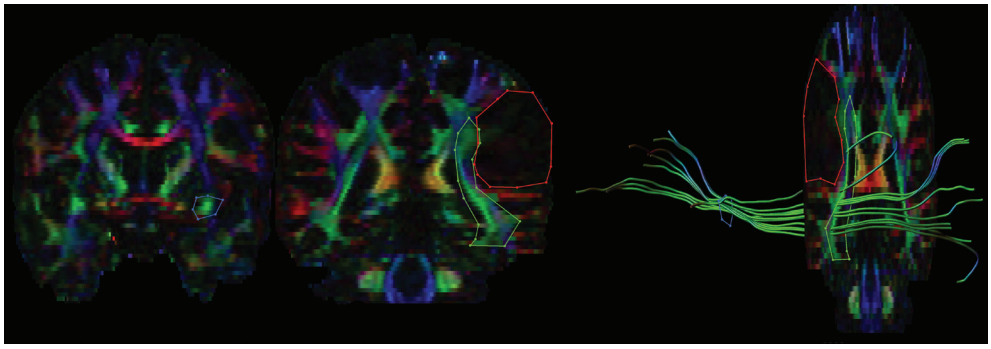


Figure 2. The first ROI is placed on the green projections in the anterior part of the frontal lobe. The second ROI is placed on the green projections in the posterior part of the occipital lobe. The tumor is delineated in red. The 3D-reconstruction of the IFOF is shown at the most right figure.

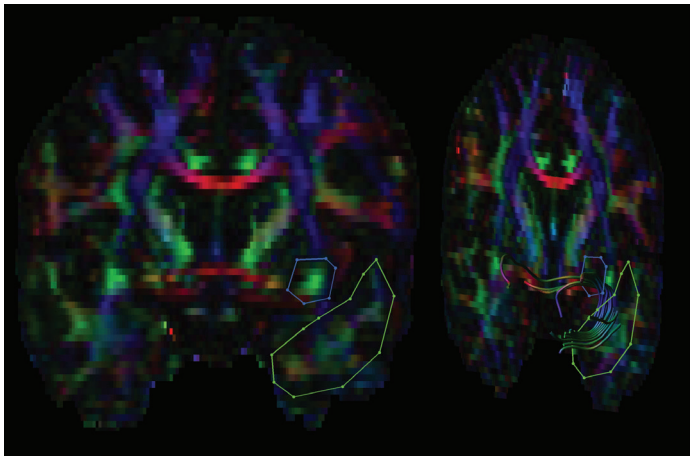


Figure 3. The first ROI is placed on the green projections in the anterior part of the temporal lobe. The second ROI is placed on the green projections in the inferior part of the frontal lobe. The tumor is delineated in red. The 3D-reconstruction of the UF is shown at the most right figure.

Table 2. Neuropsychological test protocol with z-scores and p-values per cognitive domain (standard deviations).

Neuropsychological test	N [a]	Cognitive Abilities	Description	Mean z-score (SD)	P-value
Language domain				-0.78 (0.98)	<0.0001
Aachener Aphasia Token Test (AAT TT) (14)	75	Language comprehension; severity of language disorder	Pointing to and manipulating geometric forms on verbal commands	-0.40 (1.97)	0.740
Aachener Aphasia repetition (AAT rep) (14)	72	Repetition	Repeating phonemes, words and sentences	-0.11 (0.81)	0.013
Boston Naming Test (BNT) (15)	75	Naming (word finding)	Naming 60 pictures, presented in order of word frequency and word difficulty	-1.52 (1.72)	<0.0001
Category fluency (CAT) (16)	74	Flexibility of verbal semantic thought processing; working memory	Producing words of a given category (animals and processions) within a limited time span	-1.04 (0.84)	<0.0001
Letter fluency (LET) (17)	73	Flexibility of verbal phonological thought processing; working memory	Producing words beginning with a given letter (D,A,T) within a limited time span	-0.89 (1.19)	<0.0001
Memory domain				-1.49 (1.35)	<0.0001
15 Words test (WT) imprinting	62	Verbal learning; immediate and delayed recall	Learning a list of 15 words, with 6 recall trials; 5 immediate and 1 delayed,	-1.82 (1.43)	<0.0001
15 Words test (WT) recall (18)	62			-1.22 (1.36)	<0.0001
Attention and executive functions				-0.57 (0.93)	<0.0001
Trail Making A	65	Visuomotor speed, attention	Connecting numbers placed randomly in ascending order as rapidly as possible	-0.60 (1.14)	<0.0001
Trail Making B	63	Mental flexibility, divided attention		-0.37 (1.28)	0.027
Trail Making BA (19)	62			-0.12 (1.23)	0.459
Stroop I	53	Mental speed; selective attention	Reading colour words, naming colours and naming colours of printed words, denoting another colour	-1.09 (1.39)	<0.0001
Stroop II	53			-1.12 (1.35)	<0.0001
Stroop III	51			-0.64 (1.06)	<0.0001
Stroop interference	48			-0.31 (1.09)	0.056

[a] Number of patients assessed before surgery. In some cases, the full protocol could not be applied to all patients due to several reasons. Priority was given to tests that were most relevant to the preparation of the operative procedure.

Results

DTI tractography

Successful tractography was performed in 68 patients of the AF (88%), in 76 of the IFOF (99%), and in 74 of the UF (96%). In total, 13 (17%) WM tracts could not be tracked due to very large tumor volume causing extensive brain shift. These WM tracts were not included in the analysis of association with NP tests. Mean FA was 0. 456 (SD = 0. 059) in the AF, 0.474 (SD = 0. 056) in the IFOF, and 0. 395 (SD = 0. 094) in the UF. Mean FA and MD of the three WM tracts was not significantly different ($p > 0.05$) between LGG (N = 47) and high-grade glioma (HGG, N = 30).

Patients' performance on NP

Patients had impairments in all cognitive domains (language, memory and attention/executive functions) compared to the normative group ($p<0.05$), except for three (sub) tests (Table 2). These were in the language domain (AAT Token Test) and in the attention/executive functions (TMTBA and Stroop interference) ($p > 0.05$). Post-hoc analyses showed that language impairments were significantly worse in patients with histopathological determined high grade than those with low grade tumors ($p = 0.025$). Memory and attention/executive functions were not significantly different between these groups ($p > 0.05$).

Relationship between NP and WM tracts

The relationship between the three main cognitive domains and the FA and MD values of the three WM tracts is presented in Table 3. Significant correlations were found between the AF (FA only) and the language domain ($\beta = 0.44$; $R = 0.40$; $p = 0.003$) and between the IFOF (FA only) and the memory domain ($\beta = -0.48$; $R = 0.44$; $p = 0.006$) and the attention/executive function domain ($\beta = -0.49$; $R = 0.45$; $p = 0.008$). Additional correlations that approached significance were seen between both the FA and MD of the AF, IFOF and memory and attention/executive functions domain ($0.008 < p < 0.05$). No significant correlations were found between the UF and any of the three cognitive domains assessed with the NP ($p > 0.05$).

Within the language domain, a strong significant correlation was found between the FA in the AF and the "Aachener Aphasia Repetition Test (AAT rep) " ($\beta = 0.59$; $R = 0.53$; $p < 0.0001$). The IFOF was significantly correlated with a subtest of the memory domain for verbal memory: "15 Words Test Imprinting" ($\beta = -0.55$; $R = 0.48$; $p = 0.002$) and with a subtest for selective attention within the attention/executive function domain: "STROOP I" ($\beta = -0.62$; $R = 0.50$; $p = 0.006$) (Table 4).

Table 3. Relationship between 3 neuropsychological domains (language, memory, attention/executive functions) and FA/MD values of 3 white matter tracts (AF,IFOF,UF).

	Arcuate Fasciculus		Inferior fronto-occipital fasciculus		Uncinate fasciculus	
Domain	**FA**	**MD**	**FA**	**MD**	**FA**	**MD**
LANG n=65	β = 0.44 P = 0.003[b] R=0.40	β = -0.31 P = 0.380 R=0.28	β = -0.30 P = 0.057 R=0.40	β = 0.12 P = 0.499 R=0.28	β = 0.06 P = 0.665 R=0.40	β = 0.05 P = 0.758 R=0.28
MEM N=54	β = 0.34 P = 0.029[a] R=0.44	β = -0.21 P = 0.172 R=0.37	β = -0.48 P = 0.006[b] R=0.44	β = 0.45 P = 0.019[a] R=0.37	β = 0.32 P = 0.055 R=0.44	β = -0.27 P = 0.128 R=0.37
ATT/EX N = 55	β = 0.39 P = 0.014[a] R=0.45	β = -0.33 P = 0.036a R=0.32	β = -0.49 P = 0.008[b] R=0.45	β = 0.04 P = 0.847 R=0.32	β = 0.17 P = 0.339 R=0.45	β = 0.09 P = 0.656 R=0.32

Linear regression models with NP as a dependent factor, FA/MD and age as independent factors. β = standardized coefficient R = correlation coefficient. LANG = language, MEM = memory, ATT/EX = attention/executive function, FA = fractional anisotropy, MD = mean diffusivity [a] indicates statistical significance at level of $p<0.05$ [b] indicates statistical significance at level of $p\leq0.008$, see table 4 for subtasks associations

Table 4. Association between the significant subtasks of the significant cognitive domains and fractional anisotropy in the arcuate fasciculus and inferior.

Domain	Subtask	White matter Tract
Language	AAT REP	AF, β = 0.59 P = 0.000 R=0.53
Memory	15 WT IMPR	IFOF, β = - 0.55 P = 0.002 R=0.48
Attention and executive functions	STROOP I	IFOF, β = - 0.62 P = 0.006 R=0.50

SD standard deviation, IQR Inter Quartile Range, CE contrast enhancement, FLAIR Fluid-attenuated inversion recovery.

Linear regression models with NP as a dependent factor, FA, grade and age as independent factors. β = standardized coefficient. R = correlation coefficient. AAT REP = Aachener Aphasia Repetition Test, 15 WT IMPR = 15 Words Test Imprinting AF = arcuate fasciculus, IFOF = inferior fronto-occipital fasciculus

Discussion

We performed a large-scale quantitative analysis of language WM tract microarchitecture and cognitive performance in patients with presumed LGG. Patients had deficits in all assessed cognitive domains, i.e. language, memory and attention/executive functions. A significant correlation was found between FA in the AF and the language domain, specifically with a repetition test, which is associated with phonological abilities. The FA in the IFOF was significantly correlated with deficits in verbal learning (imprinting) and selective attention. In this cohort of patients with tumors that radiologically appear to be low grade (i.e. presumed LGG), language impairments were significantly worse in patients with tumors that turned out to be high grade upon histopathological examination.

DTI measures within the language WM tracts

We found that lower FA, but not MD, was significantly correlated with language and verbal memory deficits in patients with presumed LGG. FA is a measure of directionality of molecular motion classically known to be sensitive to microstructural changes, but not specific to the type of changes. MD is a measure of magnitude of diffusion and is sensitive to cellularity, edema and necrosis [2]. Several tumor associated mechanisms are described that influence FA and MD. Tumor infiltration in WM tracts as well as edema may cause a decrease in FA and increase in MD, while compression of WM tracts due to tumor causes an increase in FA and decrease in MD [15] . Lee et al showed that purely vasogenic edema, composed purely of extracellular water (e.g. with meningiomas) causes a more relative increase in MD than a decrease in FA. On the other hand, tumor infiltrated edema - infiltration of tumor (e.g. glioma) cells in WM tracts - causes a more relative decrease in FA than an increase in MD [16]. We can therefore speculate that our findings may reflect tumor infiltration into the white matter tracts.

Arcuate fasciculus and language deficits

We demonstrated an association between microarchitectural changes of the AF and impaired language repetition abilities. The AF, also known as the dorsal pathway in the dual stream model of Hickok & Poeppel [13], is associated with mapping sound to articulatory-based representations (e.g. phonology), which explains the functional association between microstructural changes and phonological deficits observed here. Our findings also support recent results from Sierpowska et al. [29]. They reported that monitoring (non) word repetition is relevant near the AF during electro(sub)cortical stimulation for preservation of language production, in particular phonological performance [25,29].

Deficits in language repetition in combination with relatively intact comprehension, as observed in our patient group, is "classically" known as conduction aphasia in the stroke literature. Several studies in patients with different etiologies have demonstrated a link between this specific aphasia type and damage to the AF (e.g. tumor infiltration, WM

tract compression) [4]. In addition, damage to AF may be predictive for persistent language deficits [6,22] at test level, but also in relation to the quality of communication, underlining the importance of a repetition task.

Inferior fronto-occipital fasciculus and cognition

Another important pathway, apart from AF, that is reported to be related to language is the IFOF, which is part of the so-called ventral stream according to Hickok & Poeppel [13]. The ventral stream (IFOF) is involved in mapping sound on meaning (e.g. semantics, comprehension). In our study, the participants had no comprehension deficits, as measured with the Token Test, when compared with healthy people, and it is therefore not surprising that no association was found between DTI measures within the IFOF and the language domain. Within the semantic level, there were impairments in category fluency, but the multidimensional background of this test, i.e. language (lexical retrieval), semantic memory and in part also attention/executive functioning, may be responsible for the lack of a correlation. While it is possible that there were indeed no (clear) semantic deficits in our patient population, this finding may also indicate a limitation in our NP test procedure in that the applied tests for language functioning were not equivalent to the main linguistic components: semantics, syntax and phonology. In contrast with phonology, semantics and syntax were more globally examined by means of the Token Test. We have therefore developed a more elaborate test-battery for language functioning in glioma patients: the Dutch Linguistic Intraoperative Protocol (DuLIP) [7]. This is currently in the phase of evaluation.

Besides language deficits, it is suggested that damage to this WM tract can also cause other cognitive deficits. According to the model proposed by Duffau [8], the IFOF could be relevant for attention and/or (working) memory. Our study showed a correlation between FA in the IFOF and both imprinting and attention deficits, which provides additional support for previous suggestions on the involvement of the IFOF in memory.

The lack of clear cognitive deficits in other domains could be explained by neural plasticity, i.e. functional reorganization. Tumors in all our patients were presumed to be LGG (non-to-mild contrast enhancing), and in more than half of these patient's low-grade tumor was confirmed upon histopathological examination. LGG is typically associated with relatively slow tumor growth rate facilitating neural plasticity. As a result, functional reorganization could have taken place, which would lead to less severe deficits than expected from the degree of tract infiltration. More sensitive tasks to detect mild cognitive deficits in LGG patients would therefore be useful.

Tumor grade

Of note in this context is the finding that a large proportion (41. 6%) of non-to-mild enhancing glioma in this study turned out to be HGG upon histopathological examination. This could have influenced our findings if compared to a purely LGG group. However,

whether and how LGG influences FA/MD values in perilesional WM tracts differently from HGG is still under debate [19,23]. In this study, language impairments were significantly worse in histopathological established HGG than in LGG, while there was no significant difference in FA or MD values of the three WM tracts between these two groups. The former is consistent with the more aggressive nature of these tumors, the latter with their radiological characteristics. We can speculate that the higher grade of these tumors allows for less plasticity and hence worse deficits.

Limitations

There were some limitations to our study. Deterministic tractography was used to identify three WM tracts per patient by placing two ROIs per WM tract. It is important to note that the procedure of manually anatomical seed ROI placement is subject to variability within and between raters. Subjectivity depends on the location of ROI placement and the WM of interest, with a higher variability of tracking in cortical areas and higher variability within specific tracts such as the inferior longitudinal fasciculus[21,28]. This is even more relevant in the context of brain tumors, as Schonberg et al. showed that when WM tracts are displaced by excessive edema or shift due to tumor compression, the variability and subjectivity of tractography increases significantly when compared with the contralateral hemisphere [28]. We attempted to keep the inter-subject variability low and success rate high, by using a single, well-trained rater who placed ROIs systematically on clearly identifiable anatomical landmarks on DTI maps using published approaches[30] and who visually inspected each identified WM tract in three dimensions. Additionally, the exclusion of WM tracts that could not be tracked in this analysis could have introduced some selection bias, since these patients potentially have more or more severe cognitive deficits. Furthermore, one could argue that the diffusion imaging acquisition is not fully state-of-the art in the context of neuroscientific research, which now includes multiple b-values and many more gradient directions. This study was performed within the constraints of a routine clinical context, in which longer scan durations are not feasible. This may have limited our sensitivity to more subtle changes or associations, which means our negative findings need to be interpreted with care. A further reason why to interpret findings in DTI studies with care is due to the possibility that the difference in magnet strength (1.5 T vs 3.0T) can cause a difference in FA/MD value in the WM tracts[11]. This subsequently is likely to cause reduced strengths of associations between FA/MD values and cognitive outcome. However, we did not find a difference in FA/MD values between the two magnet strengths in this study and we can fairly confidently state that the associations we did find were valid and strong findings.

Finally, due to multiple testing, a stringent significance level was maintained. As a result, there is the risk of false-negative findings and relevant correlations could have been obscured. On the other hand, the risk of false-positive findings is low and correlations that did pass the stringent significance level can be considered clearly positive.

Conclusion

This clinical DTI study shows that pre-operative language deficits in repetition of speech, verbal learning and attention deficits are associated with changes in the microarchitecture of respectively the AF and the IFOF in patients with presumed LGG. This indicates that especially repetition tasks during surgery are very important and should be implemented in the NP as a standard. Verbal learning is difficult to monitor during surgery, which highlights the need for a careful pre-operative assessment of the IFOF with DTI, especially in patients with decreased performance on the 15WT. We emphasize that, especially in patients with deficits in speech repetition or verbal learning, performing extensive tumor resection will need to be balanced carefully against preserving infiltrated WM tracts (AF, IFOF). Further investigation is needed to assess the predictive value of pre-operative FA of WM tracts, tumor genetic profile, and extent of tumor resection on post-operative cognitive outcome.

References

1. Abhinav K, Yeh FC, Mansouri A, Zadeh G, Fernandez-Miranda JC: High-definition fiber tractography for the evaluation of perilesional white matter tracts in high-grade glioma surgery. **Neuro Oncol 17:**1199-1209, 2015
2. Alexander AL, Hurley SA, Samsonov AA, Adluru N, Hosseinbor AP, Mossahebi P, et al: Characterization of cerebral white matter properties using quantitative magnetic resonance imaging stains. **Brain Connect 1:**423-446, 2011
3. Bizzi A, Nava S, Ferre F, Castelli G, Aquino D, Ciaraffa F, et al: Aphasia induced by gliomas growing in the ventrolateral frontal region: assessment with diffusion MR tractography, functional MR imaging and neuropsychology. **Cortex 48:**255-272, 2012
4. Breier JI, Hasan KM, Zhang W, Men D, Papanicolaou AC: Language dysfunction after stroke and damage to white matter tracts evaluated using diffusion tensor imaging. **AJNR Am J Neuroradiol 29:**483-487, 2008
5. Catani M, Thiebaut de Schotten M: A diffusion tensor imaging tractography atlas for virtual in vivo dissections. **Cortex 44:**1105-1132, 2008
6. Caverzasi E, Hervey-Jumper SL, Jordan KM, Lobach IV, Li J, Panara V, et al: Identifying preoperative language tracts and predicting postoperative functional recovery using HARDI q-ball fiber tractography in patients with gliomas. **J Neurosurg 125:**33-45, 2016
7. De Witte E, Satoer D, Robert E, Colle H, Verheyen S, Visch-Brink E, et al: The Dutch Linguistic Intraoperative Protocol: a valid linguistic approach to awake brain surgery. **Brain Lang 140:**35-48, 2015
8. Duffau H: The anatomo-functional connectivity of language revisited. New insights provided by electrostimulation and tractography. **Neuropsychologia 46:**927-934, 2008
9. Duffau H, Capelle L, Sichez N, Denvil D, Lopes M, Sichez JP, et al: Intraoperative mapping of the subcortical language pathways using direct stimulations. An anatomo-functional study. **Brain 125:**199-214, 2002
10. Duffau H, Gatignol P, Mandonnet E, Peruzzi P, Tzourio-Mazoyer N, Capelle L: New insights into the anatomo-functional connectivity of the semantic system: a study using cortico-subcortical electrostimulations. **Brain 128:**797-810, 2005
11. Fushimi Y, Miki Y, Okada T, Yamamoto A, Mori N, Hanakawa T, et al: Fractional anisotropy and mean diffusivity: comparison between 3.0-T and 1.5-T diffusion tensor imaging with parallel imaging using histogram and region of interest analysis. **NMR Biomed 20:**743-748, 2007
12. Graetz S, De Blesser P, Willmes K: **Akense Afasie Test, Dutch edition.** Lisse: Swets & Zeitlinger, 1991
13. Hickok G, Poeppel D: Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. **Cognition 92:**67-99, 2004
14. Kaplan E, Goodglass H, Weintraub S: **Boston Naming Test.** Philadelphia-Tokyo: Lippincott, Williams and Wilkins, 2001

15. Kinoshita M, Nakada M, Okita H, Hamada J, Hayashi Y: Predictive value of fractional anisotropy of the arcuate fasciculus for the functional recovery of language after brain tumor resection: a preliminary study. **Clin Neurol Neurosurg 117:**45-50, 2014
16. Lee HY, Na DG, Song IC, Lee DH, Seo HS, Kim JH, et al: Diffusion-tensor imaging for glioma grading at 3-T magnetic resonance imaging: analysis of fractional anisotropy and mean diffusivity. **J Comput Assist Tomogr 32:**298-303, 2008
17. Leemans A, Jones DK: The B-matrix must be rotated when correcting for subject motion in DTI data. **Magn Reson Med 61:**1336-1349, 2009
18. Lezak MD: **Neuropsychological Assessment.** New York: Oxford University Press, 2004
19. Lu S, Ahn D, Johnson G, Law M, Zagzag D, Grossman RI: Diffusion-tensor MR imaging of intracranial neoplasia and associated peritumoral edema: introduction of the tumor infiltration index. **Radiology 232:**221-228, 2004
20. Luteijn F, Barelds DPF: **Groninger Intelligentie Test II (GIT II).** Amsterdam: Pearson, 2004
21. Makris N, Kennedy DN, McInerney S, Sorensen AG, Wang R, Caviness VS, Jr., et al: Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. **Cereb Cortex 15:**854-869, 2005
22. Marchina S, Zhu LL, Norton A, Zipse L, Wan CY, Schlaug G: Impairment of speech production predicted by lesion load of the left arcuate fasciculus. **Stroke 42:**2251-2256, 2011
23. Miloushev VZ, Chow DS, Filippi CG: Meta-analysis of diffusion metrics for the prediction of tumor grade in gliomas. **AJNR Am J Neuroradiol 36:**302-308, 2015
24. Papagno C, Casarotti A, Comi A, Gallucci M, Riva M, Bello L: Measuring clinical outcomes in neuro-oncology. A battery to evaluate low-grade gliomas (LGG). **J Neurooncol 108:**269-275, 2012
25. Sarubbo S, De Benedictis A, Merler S, Mandonnet E, Balbi S, Granieri E, et al: Towards a functional atlas of human white matter. **Hum Brain Mapp 36:**3117-3136, 2015
26. Satoer D, Visch-Brink E, Smits M, Kloet A, Looman C, Dirven C, et al: Long-term evaluation of cognition after glioma surgery in eloquent areas. **J Neurooncol 116:**153-160, 2014
27. Schmand B, Groenink SC, Van den Dungen M: Letterfluency: psychometrische eigenschappen en Nederlandse normen. **Tijdschrift voor Gerontologie en Geriatrie 39:**65-77, 2008
28. Schonberg T, Pianka P, Hendler T, Pasternak O, Assaf Y: Characterization of displaced white matter by brain tumors using combined DTI and fMRI. **Neuroimage 30:**1100-1111, 2006
29. Sierpowska J, Gabarros A, Fernandez-Coello A, Camins A, Castaner S, Juncadella M, et al: Words are not enough: nonword repetition as an indicator of arcuate fasciculus integrity during brain tumor resection. **J Neurosurg:**1-11, 2016
30. Smits M, Jiskoot LC, Papma JM: White matter tracts of speech and language. **Semin Ultrasound CT MR 35:**504-516, 2014
31. Talacchi A, d'Avella D, Denaro L, Santini B, Meneghelli P, Savazzi S, et al: Cognitive outcome as part and parcel of clinical outcome in brain tumor surgery. **J Neurooncol 108:**327-332, 2012

32. Thiebaut de Schotten M, Ffytche DH, Bizzi A, Dell'Acqua F, Allin M, Walshe M, et al: Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. **Neuroimage 54:**49-59, 2011
33. Van der Elst W, van Boxtel MP, van Breukelen GJ, Jolles J: Rey's verbal learning test: normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. **J Int Neuropsychol Soc 11:**290-302, 2005

PART II

Image-guided glioblastoma surgery

Chapter 6

Intraoperative ultrasound guided surgery and the extent of glioblastoma resection: a randomized, controlled trial

Fatih Incekara, Marion Smits, Linda Dirven, Eelke Bos, Rutger Balvers, Iain Haitsma, Joost Schouten, A.J.P.E. Vincent

Submitted

Abstract

Background

Intraoperative MRI and 5-aminolaevulinic acid guided surgery is useful to maximize the extent of glioblastoma resection. Intraoperative ultrasound is used as a time-and cost effective alternative, but its value has never been assessed in a trial. The goal of this randomized controlled trial was to assess the value of intraoperative ultrasound guided surgery on the extent of glioblastoma resection.

Methods

In this randomized controlled trial, patients of 18 years or older with a newly diagnosed presumed glioblastoma, deemed totally resectable, presenting at the Erasmus MC (Rotterdam, The Netherlands) were enrolled and randomized (1:1) into intraoperative ultrasound guided surgery or standard surgery with neuronavigation alone. The primary outcome of this study was complete contrast-enhancing tumor resection, assessed quantitatively by a blinded neuroradiologist on pre- and post-operative MRI scans. This trial is registered with ClinicalTrials.gov (NCT03531333).

Results

We enrolled 50 patients between November 1, 2016 and October 30, 2019. Analysis was done in 23 of 25 (92%) patients in the intraoperative ultrasound group and 24 of 25 (96%) patients in the standard surgery group. Eight (35%) of 23 patients in the intraoperative ultrasound group and two (8%) of 24 patients in the standard surgery group underwent complete resection ($p=0.036$). Baseline characteristics, neurological outcome, functional performance, quality of life, complication rates, overall survival and progress-free survival did not differ between treatment groups ($p>0.05$).

Conclusion

Intraoperative ultrasound enables complete resection more often than standard surgery without harming patients and can be considered to maximize the extent of glioblastoma resection during surgery.

Introduction

Patients with glioblastoma have a poor prognosis with a median overall survival of 15 months, despite surgical resection with concomitant and adjuvant chemoradiotherapy.[1] Complete resection of contrast-enhancing tumor on T1-weighted post-contrast MRI has consistently been associated with longer overall survival.[2] This association has recently also been confirmed in molecular subgroups of glioblastoma.[3] The same study shows that resection beyond the borders of contrast enhancement is associated with improved overall survival in patients with glioblastoma.

It is shown that intraoperative technologies, specifically 5-aminolevulinic acid or intraoperative MRI guided surgery, are useful to maximize tumor resection during glioblastoma surgery.[4-6] Although intraoperative MRI has been associated with higher rates of complete glioblastoma resection, its use is expensive and time-consuming.[6] Intraoperative ultrasound guidance is used during glioblastoma surgery as a time- and cost-effective intraoperative imaging alternative.[7] As Jenkinson et al. showed in a Cochrane review however, its value to maximize tumor resection has never been assessed in a randomized controlled trial.[4]

We therefore initiated the first randomized controlled trial assessing the value of intraoperative ultrasound guided surgery on the extent of glioblastoma resection.

Methods

In this randomized controlled trial, patients of 18 years or older with a newly diagnosed, contrast-enhancing presumed glioblastoma, deemed totally resectable, presented at the Erasmus MC (Rotterdam, The Netherlands) were enrolled. Exclusion criteria were tumors located in the basal ganglia, cerebellum, brain stem or crossing the midline thereby prohibiting complete resection; multifocal tumors; patients with a Karnofsky performance status < 60 or with pre-existing neurological deficits (e.g. aphasia, hemiparesis). The study was approved by the Medical Ethical Committee of Erasmus MC (MEC-2015-46). All patients gave written informed consent prior to participation. This trial was reported following the CONSORT guidelines and registered with ClinicalTrials.gov (NCT03531333).

Randomization and Intervention

We randomly assigned patients (1:1) into intraoperative ultrasound guided surgery (intervention) or standard surgery with neuronavigation alone, without ultrasound guidance (control). Randomization was done via www.sealedenvelope.com with use of random computer-generated blocks of four by a research assistant who was not otherwise involved with this study. Neurosurgeons and patients were not blinded for treatment allocation. The primary outcome assessor, an independent neuroradiologist, was blinded for treatment group allocation.

Intraoperative ultrasound guidance was performed with an ultrasound system (BK Medical Flex Focus 800 with transducer 8862) alone or integrated with a neuronavigation system (Brainlab, Munich, Germany). Intraoperative ultrasound was used before opening of the dura to locate the tumor, during tumor resection and to locate any residual tumor in the surgical cavity. Resection was continued until no residual tumor suspected, hyperechoic lesion as seen on ultrasound images was observed in the surgical cavity, or until further resection was deemed unsafe.

Standard surgery was performed with conventional neurosurgical techniques, such as neuronavigation system, cavitational ultrasonic surgical aspiration and surgical microscope. After wound closure, surgeons were asked in both treatment groups to estimate whether complete tumor resection was achieved (yes or no). Surgery time was measured from skin incision to wound closure. Standard adjuvant chemo-and or radiotherapy and clinical follow-up with periodic MRI scans were followed for patients in both groups.[8]

Outcome Measures

The primary outcome of this study was complete resection of contrast-enhancing tumor on early postoperative MRI. All patients underwent 1.5T or 3T MRI scanning with and without gadolinium-based contrast agent one day before surgery and within 48 hours after surgery. One blinded, independent, highly experienced neuroradiologist assessed the tumor localization and extent of tumor resection by volumetrically measuring initial and residual contrast-enhancing tumor volumes. First, pre- and post-operative T1-weighted contrast scans were loaded into Brainlab Elements. Using the SmartBrush tool, semi-automatic tumor assessment of all tumor involved contrast enhancement on preoperative scans and on post-operative scans (excluding small vessels or blood in the surgical cavity) was performed. Tumor localization in terms of eloquence was rated following the Sawaya classification (grade 1: non-eloquent, grade 2: near eloquent, grade 3; eloquent).[9] Complete resection was defined as ≥ 99% resection of contrast-enhancing tumor volume.

Secondary outcomes were: extent of tumor resection (%); neurological status on the National Institutes of Health Stroke Scale (NIHSS) within one week after surgery; functional status on Karnofsky performance scale seven weeks, three months and six months after surgery; change over time in health-related quality of life (EORTC QLQ-C-30[10] and QLQ-BN20[11,12] questionnaire) from baseline up to six months after surgery; complication rates; overall survival and progression free survival. EORTC scoring procedures were followed to calculate scale scores.[13] Three QLQ-C30 scales (global health, physical functioning, cognitive functioning) and two QLQBN-20 scales (motor dysfunction and communication deficits) were preselected for analysis. A change over time of ≥10 points were classified as clinically meaningful changes.[14] Complications were classified according to the US National Cancer Institute common toxicity criteria (CTCAE, version 4.0). Overall survival

was defined as time from surgery to death and progression-free survival was defined as time from surgery till clinical or radiological progression following the RANO criteria.[15]

Statistical Analysis

Sample size calculation was performed for the primary outcome based on retrospective studies on intraoperative ultrasound and standard surgery as described in the trial protocol (Supplemental Material). Based on complete resection rates of the conventional treatment arm as reported by Stummer et al. we estimated that in the standard surgery group 36% of patients would have complete tumor resection.[5] With an estimated effect size of 40% increase of complete resection proportion, power of 80% and significance level at 0.05, we calculated that each group had to include 23 patients. To account for the possibility of drop-out or missing data, we increased the sample size to 25 patients per treatment arm and a total of 50 patients.

Statistical analyses were performed with SPSS 25.0 statistical software (IBM Corp.). Descriptive statistics were tested between treatment arms with the Chi Squared test or Fisher Exact test in case of categorical variables and with the Mann-Whitney U test in case of continuous non-normal distributed data. Survival data were compared between treatment groups with log rank tests and Kaplan Meier estimates and analyzed with multivariable Cox proportional-hazards models. Linear Mixed Models were used to compare health-related quality of life scores over time between treatment arms.

Results

We enrolled 50 patients between November 1, 2016 and October 30, 2019. Two patients who were diagnosed with metastases after surgery in the ultrasound group and one patient who received a biopsy instead of surgery in the control group were excluded from all further analyses (Figure 1). Patient and tumor baseline characteristics did not differ between treatment groups (Table 1).

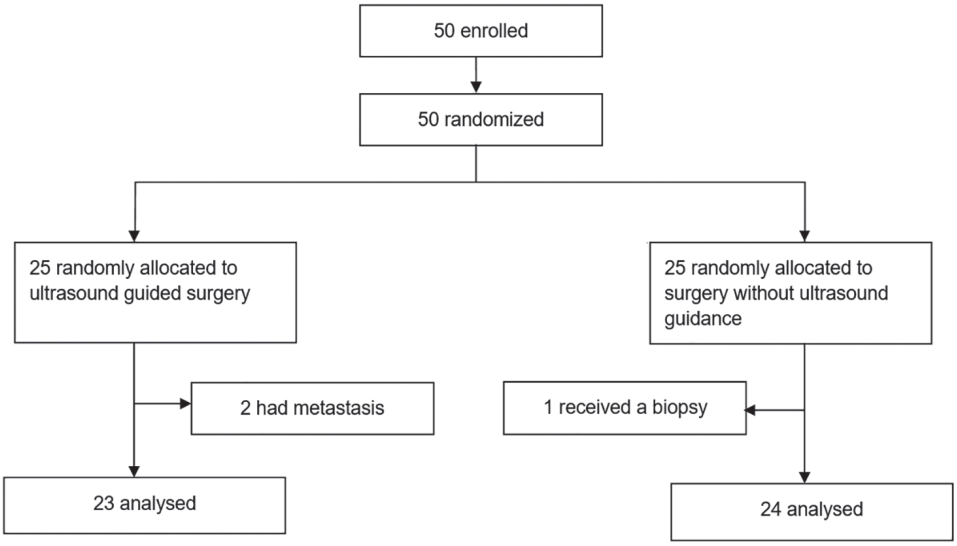


Figure 1. Flow chart.

Table 1. Baseline characteristics.

	Intraoperative ultrasound (n=23)	**Standard surgery (n=24)**
Age, median years (IQR)	62 (54-71)	64 (57-70)
Sex		
Male	14 (61%)	14 (58%)
Female	9 (39%)	10 (42%)
KPS, median (IQR)	90 (80-100)	90 (80-100)
Tumor localization*		
Non-eloquent	8 (35%)	8 (33%)
Near eloquent	6 (26%)	6 (25%)
Eloquent	9 (39%)	10 (42%)
Tumor volume, median cm^3 (IQR)	38.6 (16.9-60.1)	32.3 (17.2-44.6)
NIHSS, median (IQR)	1 (0-1)	0 (0-2)
Quality of life, mean (SD)**		
Global health status	75 (24)	77 (17)
Physical functioning	88 (15)	91 (16)
Cognitive functioning	88 (16)	85 (21)
Motor dysfunction	12 (18)	10 (21)
Communication deficit	17 (24)	9 (14)

Data are No. (%), unless stated otherwise. * Sawaya Grading System ** For global health status, physical functioning and cognitive functioning, a higher score represents better functioning. For motor dysfunction and communication deficit, a higher score represents more problems. KPS Karnofsky performance status, IDH isocitrate dehydrogenase, MGMT methylguanine DNA methyltransferase, NIHSS National Institutes of Health stroke score.

Eight (35%) of 23 patients in the intraoperative ultrasound group and two (8%) of 24 patients in the standard surgery group had complete resection of contrast-enhancing tumor (odds ratio 5.9 (95% CI 1.1-31.6), p=0.036; proportion difference 27% (95% CI, 2.8-47.7), p=0.024). Median extent of resection was 97% (IQR 89-100) with intraoperative ultrasound and 95% (IQR 79-98) with standard surgery (p=0.151, Table 2). Median residual tumor volume was 0.9 cm^3 (IQR 0.2-3.4) with intraoperative ultrasound and 1.4 cm^3 (IQR 0.7-6.4) with standard surgery (p=0.205). Patient outcome of both treatment groups are presented in Table 3.

Table 2. Surgery outcome.

	Intraoperative ultrasound (n=23)	**Standard surgery (n=24)**	**p value**
Resection			**0.036**
Complete	8 (35%)	2 (8%)	
Incomplete	15 (65%)	22 (92%)	
Extent of resection, median (IQR), %	97 (89-100)	95 (79-98)	0.151
Residual tumor volume, median (IQR), cm^3	0.9 (0.2-3.4)	1.4 (0.7-6.4)	0.205
Surgery time, median (IQR), minutes	177 (135-255)	179 (146-227)	0.907
Blood loss, median (IQR), ml	150 (0-400)	125 (58-200)	0.729

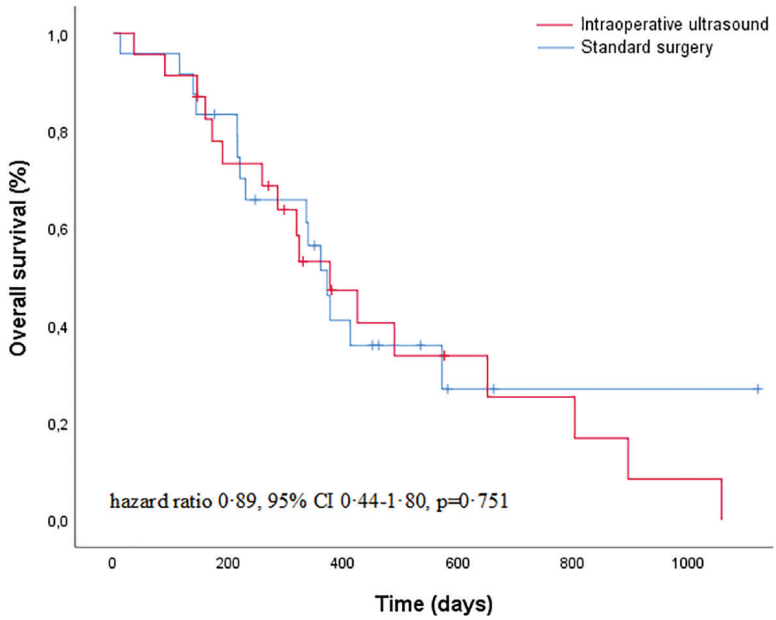
Data are n or n (%), unless stated otherwise.

Intraoperative ultrasound was used four times (range two to nine) on average per surgery. In the operating room, surgeons estimated that complete tumor resection was achieved in 15 (65%) of 23 patients when intraoperative ultrasound was used and in 17 (71%) of 23 patients without the use of intraoperative ultrasound (p=0.680). However, cases in which complete resection was thought to be achieved corresponded with radiological complete resection in only two (11.8%) of 17 in the standard surgery group and in seven (46.7%) of 15 patients in the intraoperative ultrasound group (proportion difference 34.9%, 95% CI 3.5-59.6, p=0.031; odds ratio 6.6, 95% CI 1.1-39.3, p=0.049). Median surgery time with intraoperative ultrasound guided surgery (177 minutes, IQR 135-255) was comparable to standard surgery (179 minutes, IQR 146-227, p=0.907). Kaplan Meier estimates showed that median overall survival was 377 days (95% CI 247-507) in the intraoperative ultrasound group and 372 days (95% CI 320-424) in the standard surgery group (hazard ratio 0.89, 95% CI 0.44-1.80, p=0.751, Figure 2). Median progression-free survival was 277 days (95% CI 107-347) in the intraoperative ultrasound group and 233 days (95% CI 153-313) in the standard surgery group (hazard ratio 0.97, 95% CI 0.49-1.95, p=0.937). Multivariable Cox proportional-hazards model analysis adjusted for potential prognostic factors (age, MGMT promotor methylation, tumor volume and localization) indicated that an increase in extent of resection percentage was significantly associated with overall survival (hazard ratio 0.97, 95% CI 0.95-0.99, p=0.037).

Table 3. Patient outcome.

	Intraoperative ultrasound (n=23)	Standard surgery (n=24)	p value
IDH mutation			**0.494**
Mutated	0 (0%)	0 (0%)	
Wildtype	19 (83%)	17 (71%)	
Unknown	4 (17%)	7 (29%)	
MGMT promotor methylation			**0.347**
Methylated	6 (26%)	7 (29%)	
Unmethylated	13 (57%)	9 (38%)	
Unknown	4 (17%)	8 (33%)	
Adjuvant therapy			**0.148**
None	3 (13%)	2 (8%)	
Chemo or radiotherapy	3 (13%)	0 (0%)	
Chemoradiation*	17 (74%)	22 (92%)	
NIHSS post-operative, median (IQR)	0 (0-2)	0 (0-2)	0.825
KPS after surgery, median (IQR)			
Seven weeks	90 (90-100)	90 (80-100)	0.412
Three months	90 (80-100)	90 (70-100)	0.540
Six months	90 (70-90)	70 (60-90)	0.228
Quality of life change, baseline vs. six months**			
Global health status	-2 (35)	-14 (28)	0.344
Physical functioning	-8 (31)	-13 (18)	0.267
Cognitive functioning	-11 (32)	-2 (30)	0.893
Motor dysfunction	2 (21)	5 (20)	0.893
Communication deficit	1 (26)	-6 (22)	0.609
Overall survival, median (95% CI), days	377 (247-507)	372 (320-424)	0.751
Progression-free survival, median (95% CI), days	227 (107-347)	233 (153-313)	0.937

Data are n or n (%), unless stated otherwise. * Stupp protocol ** A change of ≥10 points is considered to be clinically relevant. KPS Karnofsky performance status, IDH isocitrate dehydrogenase, MGMT methylguanine DNA methyltransferase, NIHSS National Institutes of Health stroke score.



Numbers at risk						
Intraoperative ultrasound	23	16	7	4	3	1
Standard surgery	24	19	8	2	1	1

Figure 2. Kaplan-Meier curve for overall survival by treatment group.

Median KPS seven weeks and three months after surgery was 90 (IQR 70/80-100) in both treatment groups. Six months after surgery, KPS was 60 or below in three patients (17%) who underwent intraoperative ultrasound surgery and in seven patients (37%) who underwent standard surgery (p=0.269). Neurological outcome as measured using the NIHSS scale within one week after surgery did not significantly differ between treatment groups (NIHSS 0 (IQR 0-2), p=0.825). In the intraoperative ultrasound vs. standard surgery groups, 16 (70%) respectively 19 (79%) patients had the same neurological status on the NIHSS scale after surgery as before surgery and five (22%) respectively three (13%) patients had neurological improvement. Four (9%) of all 47 patients had new or worsened neurological deficits: two (8%) patients who underwent intraoperative ultrasound guided surgery (one patient with hemiparesis and one with delirium and superior sagittal sinus thrombosis) and two patients (8%) who underwent standard surgery (one patient with aphasia and one with postoperative hemorrhage). Characteristics of these patients are presented in more detail in Table 4. Frequency of new or worsened neurological deficits did not significantly differ between treatment groups (p=0.591). Compliance with health-related quality of life assessments was 100% at baseline and 69% at 6-months. Preoperative health-related quality of life scores were comparable between treatment arms (Table 1). Mean change in health-related quality of life scores from baseline to the

six-month assessment was not significantly different between treatment arms (p>0.05, Table 3). The linear mixed model analyses showed that for none of the preselected health-related quality of life scales the overall mean difference over six months' time, on a group level, was significantly different or clinically relevant between treatment arms: global health=-3.9 (p=0.526), physical functioning=-0.4 (p=0.946), cognitive functioning=-3.5 (p=0.554), motor dysfunction=0.9 (p=0.875) and communication deficit=-9.3 (p=0.114).

Table 4. Details of patients with complications.

ID	Sex	Age	Treatment group	NIHSS pre-op	NIHSS post-op	CTCAE	Complication	Treatment	KPS 7WK-3M-6M	Survival (days)
1	Male	75	Standard surgery	1	5	2	Aphasia	None	80-80-70	377
7	Male	65	Standard surgery	0	2	4	Postoperative hemorrhage	Emergency craniotomy	80-80-60	361
30	Male	59	Ultrasound	1	9	2	Delirium, sagittal sinus thrombosis	Haldol, Fraxiparine	n/a	36
35	Male	43	Ultrasound	0	14	3	Hemiparesis, central facial palsy	None	50-40-n/a	172

NIHSS National Institutes of Health Stroke Scale, CTCAE Common Terminology Criteria for Adverse Events grading v.4.0, KPS Karnofsky Performance Status, n/a not applicable.

Discussion

This is the first randomized controlled trial that assessed the value of intraoperative ultrasound guided surgery on the extent of glioblastoma resection. Our trial showed that intraoperative ultrasound guided surgery enables complete contrast-enhancing tumor resections more often than standard surgery, without harming patients in terms of neurological outcome, functional performance or health-related quality of life.

Complete resection of contrast-enhancing tumor during glioblastoma surgery has consistently been associated with longer overall survival.[2] This association was recently confirmed across different molecular subgroups of glioblastoma based on WHO 2016 classification including IDH mutation status and MGMT methylation status.[3] In addition, studies have shown that resection of non-contrast-enhancing tumor beyond the borders of contrast enhancement is associated with longer overall survival.[3,16,17] It is shown that 5-aminolevulinic acid and intraoperative MRI guided surgery improves the extent of glioblastoma resection.[4-6,18-20] Stummer et al. showed in a randomized controlled trial that complete contrast-enhancing tumor resection is achieved more often with 5-aminolevulinic acid guided surgery (65%) than with standard surgery (36%).[5] It is also shown in a randomized controlled trial that with a low field (0.15 T), intraoperative MRI more frequent complete resections of contrast-enhancing glioblastoma was achieved (96%) than with standard surgery (68%). As Senft and colleagues noted however, an

intraoperative MRI system is expensive and prolongs surgery time with approximately one hour.[6]

An alternative potentially cost- and time-effective technology that is used to acquire real-time imaging and apply brain shift correction during neuro-oncological surgery is intraoperative ultrasound guidance.[7] Retrospective studies have suggested that intraoperative ultrasound may increase the extent of tumor resection during surgery.[21-23] These studies however, included different glioma subtypes and held different definitions of gross total resection, thereby introducing some degree of selection and confounding biases. No randomized controlled trial was performed to date to assess the value of intraoperative ultrasound to maximize the extent of glioblastoma resection.[4]

When compared to intraoperative MRI, intraoperative ultrasound has two advantages; it is less expensive and, as shown in our trial, it does not prolong surgery time. Surgeons could rapidly control for residual tumor in the resection cavity multiple times during surgery without prolonging surgery time. The interpretation of intraoperative ultrasound images might be more challenging than intraoperative MRI images, however, the integration of an intraoperative ultrasound systems with a standard neuronavigation system (as used in our trial) enables surgeons to overlay intraoperative ultrasound images on navigational preoperative MRI scans, which may facilitate the interpretation of ultrasound images and consequently the accuracy of complete tumor resection estimation. Importantly, we observed that when intraoperative ultrasound was used, surgeons were able to estimate complete tumor resection in the operating room significantly more accurately than with standard surgery without ultrasound guidance, as confirmed on post-operative MRI.

Complete tumor resection as a primary outcome has some aspects that need careful consideration. Several definitions of complete tumor resection exist across studies, both qualitatively as quantitatively.[2,18] Studies defined complete tumor resection as no residual contrast-enhancing tumor on a post-operative MRI scan[17,24], which is a relatively stringent definition (if quantitatively assessed) and may result in false positive assessment of the presence of residual tumor due to non-specific contrast enhancement such as ischemia, small vessels, a non-specific tissue response, or by T1-hyperintense blood in the surgical cavity that is incorrectly interpreted as enhancement. In our trial, this was mitigated by overlaying the identically acquired and registered pre- and post-contrast T1w sequences to exclude any T1-hyperintense areas from the residual tumor delineation. To take interpretation varieties into account, some studies defined complete tumor resection as contrast-enhancing residual tumor smaller than 0.175 cm^3 following Stummer et al.[5,6] The rationale behind this cut-off point was that it was the volume of one single voxel and the experience in assessing residual tumor volumes of a neuroradiologist in their hospital. Other studies have used extent of resection cut-off percentages, such as 95%, 97% or 98% resection of all contrast enhancement.[20,25-28] In relation to this, it is known that residual tumor assessment of glioblastoma has a low interobserver agreement, introducing some degree of subjectivity when distinguishing contrast-enhancing residual tumor from non-specific contrast enhancement.[29]

In this trial, complete tumor resection was defined as more than 99% resection of contrast-enhancing tumor volume, accepting residual contrast-enhancing volume smaller than one percent to account for the non-tumor related post-surgical reactive enhancement amongst others, which is present even on early (within 48h) post-operative MRI scans.[30] Even then, our complete resection proportion is lower than that reported in conventional treatment arms of previous trials, however the median extent of resection in both groups were high (97% in intraoperative ultrasound vs. in 95% standard surgery).[5,6] This indicates that the low proportion of complete resection could partially be explained by a possible stringent interpretation of small contrast-enhancing voxels in the surgical cavity rather than surgical performance, as described earlier. This may have led to false positive interpretation of residual contrast enhanced tumor (i.e. false negative complete resection outcomes) in our trial.

Importantly, we included only glioblastoma that were deemed complete resectable prior to surgery, which could partially explain the high median extent of resection percentages in both treatment groups (97% with ultrasound guidance and 95% with standard surgery). Our hypothesis in this trial was that in these totally resectable deemed glioblastoma, intraoperative ultrasound would be useful to resect the last small tumor portions and thus to actually achieve complete resection more often. As mentioned earlier, since only high resection cut-off percentage (e.g. >97% and >98%)[25,28] are associated with survival benefit, we chose complete resection, rather than resection percentage as primary outcome.

A limitation of this trial is that it was not double-blinded, however, complete resection of contrast-enhancing tumor, our primary outcome, was assessed by an independent, blinded neuroradiologist. The standard treatment arm in in this trial was performed with conventional surgical techniques such as neuronavigation, ultrasonic aspirator and surgical microscope. Senft et al. suggested that conventional treatment arms in trials studying extent of resection in brain tumors should at least be performed with either intraoperative MRI or 5-aminolevulinic acid guidance.[6] However, the use of these techniques during glioblastoma surgery is not yet adopted as standard surgical treatment in the Netherlands. As in the 5-aminolevulinic acid and intraoperative MRI trials, our trial did not show any overall survival benefit for patients who underwent intraoperative ultrasound.[5,6] It should be noted however, that our trial did not aim and was not designed or powered to show differences in overall survival, as in the intraoperative MRI and 5-ALA trials. Although complete glioblastoma resection is associated with survival benefit, future trials on image guidance with a suited design and larger sample size are still needed to show any potential clinical benefit directly in the trial itself for patients with glioblastoma.

Conclusion

This randomized controlled trial showed that intraoperative ultrasound guided surgery enables complete contrast-enhancing tumor resections more often than standard surgery, without harming patients in terms of neurological outcome, functional performance or health-related quality of life. Intraoperative ultrasound is a safe and useful intraoperative imaging alternative and, just as intraoperative MRI or 5-aminolevulinic acid guided surgery, can be considered to maximize the extent of contrast-enhancing glioblastoma resection.

References

1. Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* 2009;10(5):459-466.
2. Brown TJ, Brennan MC, Li M, et al. Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis. *JAMA Oncol.* 2016;2(11):1460-1469.
3. Molinaro AM, Hervey-Jumper S, Morshed RA, et al. Association of Maximal Extent of Resection of Contrast-Enhanced and Non-Contrast-Enhanced Tumor With Survival Within Molecular Subgroups of Patients With Newly Diagnosed Glioblastoma. *JAMA Oncol.* 2020:10.1001/jamaoncol.2019.6143.
4. Jenkinson MD, Barone DG, Bryant A, et al. Intraoperative imaging technology to maximise extent of resection for glioma Review. *Cochrane Database Syst Rev.* 2018;1:CD012788.
5. Stummer W, Pichlmeier U, Meinel T, Wiestler OD. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. *Lancet Oncol.* 2006.
6. Senft C, Bink A, Franz K, Vatter H, Gasser T, Seifert V. Intraoperative MRI guidance and extent of resection in glioma surgery: a randomised, controlled trial. *Lancet Oncol.* 2011;12(11):997-1003.
7. Unsgaard G, Ommedal S, Muller T, Gronningsaeter A, Nagelhus Hernes TA. Neuronavigation by intraoperative three-dimensional ultrasound: initial experience during brain tumor resection. *Neurosurgery.* 2002;50(4):804-812; discussion 812.
8. Stupp R, Mason WP, Van Den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New Engl J Med.* 2005;352(10):987-996.
9. Sawaya R, Hammoud M, Schoppa D, et al. Neurosurgical outcomes in a modern series of 400 craniotomies for treatment of parenchymal tumors. *Neurosurgery.* 1998;42(5):1044-1055; discussion 1055-1046.
10. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst.* 1993;85(5):365-376.

11. Osoba D, Aaronson NK, Muller M, et al. The development and psychometric validation of a brain cancer quality-of-life questionnaire for use in combination with general cancer-specific questionnaires. *Qual Life Res.* 1996;5(1):139-150.
12. Taphoorn MJ, Claassens L, Aaronson NK, et al. An international validation study of the EORTC brain cancer module (EORTC QLQ-BN20) for assessing health-related quality of life and symptoms in brain cancer patients. *Eur J Cancer.* 2010;46(6):1033-1040.
13. Fayers PM AN, Bjordal K, Groenvold M, Curran D, Bottomley A. . The EORTC QLQ-C30 Scoring Manual. 3rd ed. ed. *Brussels, Belgium: European Organisation for Research and Treatment of Cancer.* 2001.
14. Taphoorn MJ, Stupp R, Coens C, et al. Health-related quality of life in patients with glioblastoma: a randomised controlled trial. *Lancet Oncol.* 2005;6(12):937-944.
15. Wen PY, Macdonald DR, Reardon DA, et al. Updated response assessment criteria for high-grade gliomas: Response assessment in neuro-oncology working group. *J Clin Oncol.* 2010;28(11):1963-1972.
16. Pessina F, Navarria P, Cozzi L, et al. Maximize surgical resection beyond contrast-enhancing boundaries in newly diagnosed glioblastoma multiforme: is it useful and safe? A single institution retrospective experience. *J Neuro-Oncol.* 2017;135(1):129-139.
17. Li YM, Suki D, Hess K, Sawaya R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? *J Neurosurg.* 2016;124(4):977-988.
18. Kubben PL, ter Meulen KJ, Schijns OEMG, ter Laak-Poort MP, van Overbeeke JJ, van Santbrink H. Intraoperative MRI-guided resection of glioblastoma multiforme: A systematic review. *Lancet Oncol.* 2011;12(11):1062-1070.
19. Ferraro N, Barbarite E, Albert TR, et al. The role of 5-aminolevulinic acid in brain tumor surgery: a systematic review. *Neurosurg Rev.* 2016;39(4):545-555.
20. Hatiboglu MA, Weinberg JS, Suki D, et al. Impact of Intraoperative High-Field Magnetic Resonance Imaging Guidance on Glioma Surgery: A Prospective Volumetric Analysis. *Neurosurgery.* 2009;64(6):1073-1081.
21. Mahboob S, McPhillips R, Qiu Z, et al. Intraoperative Ultrasound-Guided Resection of Gliomas: A Meta-Analysis and Review of the Literature. *World Neurosurg.* 2016;92:255-263.
22. Solheim O, Selbekk T, Jakola AS, Unsgård G. Ultrasound-guided operations in unselected high-grade gliomas-overall results, impact of image quality and patient selection. *Acta Neurochir.* 2010;152(11):1873-1886.
23. Moiraghi A, Prada F, Delaidelli A, et al. Navigated Intraoperative 2-Dimensional Ultrasound in High-Grade Glioma Surgery: Impact on Extent of Resection and Patient Outcome. *Oper Neurosurg (Hagerstown).* 2020;18(4):363-373.
24. Gessler F, Bernstock JD, Braczynski A, et al. Surgery for Glioblastoma in Light of Molecular Markers: Impact of Resection and MGMT Promoter Methylation in Newly Diagnosed IDH-1 Wild-Type Glioblastomas. 2018.

25. Lacroix M, Abi-Said D, Fourney DR, et al. A multivariate analysis of 416 patients with glioblastoma multiforme: Prognosis, extent of resection, and survival. *J Neurosurg.* 2001;95(2):190-198.
26. Chaichana KL, Jusue-Torres I, Navarro-Ramirez R, et al. Establishing percent resection and residual volume thresholds affecting survival and recurrence for patients with newly diagnosed intracranial glioblastoma. *Neuro Oncol.* 2014;16(1):113-122.
27. Grabowski MM, Recinos PF, Nowacki AS, et al. Residual tumor volume versus extent of resection: Predictors of survival after surgery for glioblastoma. *J Neurosurg.* 2014;121(5):1115-1123.
28. Incekara F, Smits M, van der Voort SR, Dubbink HJ, Atmodimedjo N, Kros, JM et al. The Association Between the Extent of Glioblastoma Resection and Survival in Light of MGMT Promoter Methylation in 326 Patients With Newly Diagnosed IDH-Wildtype Glioblastoma. Front. Oncol., 10 July 2020 | https://doi.org/10.3389/fonc.2020.01087
29. Kubben PL, Postma AA, Kessels AG, van Overbeeke JJ, van Santbrink H. Intraobserver and interobserver agreement in volumetric assessment of glioblastoma multiforme resection. *Neurosurgery.* 2010;67(5):1329-1334.
30. Bette S, Gempt J, Huber T, et al. Patterns and Time Dependence of Unspecific Enhancement in Postoperative Magnetic Resonance Imaging after Glioblastoma Resection. *World Neurosurg.* 2016;90:440-447.

Supplementary Materials

Ultrasound Trial Protocol (ClinicalTrials.gov: NCT03531333)

SUMMARY

The main goal of high grade glioma (HGG) surgery is to achieve gross total resection (GTR) without causing new neurological deficits[1-8]. Intraoperative navigated high resolution ultrasound (US) is a promising new tool to acquire real-time intraoperative images to localize and to resect gliomas[9-12]. The purpose of this study is to investigate the effectivity of intraoperative navigated US in achieving GTR in patients with HGG, measure influence on quality of life and cost effectiveness.

Hypothesis: Ultrasound guided high grade glioma surgery succeeds gross total resection more frequently and improves quality of life and survival of time when compared with surgery without ultrasound guidance.

Objective:

The aim of this study was to investigate whether intraoperative guided surgery leads to a higher rate of GTR, when compared with standard non-ultrasound guided surgery.

Study design:

The US-GLIOMA study is a randomized controlled trial with blinded primary outcome measure.

Study population:

Patients with newly diagnosed contrast enhancing presumed high grade glioma on first MRI scan.

Intervention (if applicable):

The study consists of two treatment arms: non-ultrasound guided glioma resection (conventional treatment) versus ultrasound guided glioma resection (intervention) .

Main study parameters/endpoints:

- Extent of resection (Gross-total (100%) resection or sub-total (<100%) resection)
- Extent of resection (%)
- Neurological outcome (Karnofsky Performance Status)
- Quality of Life (EORTC QLQ-C30 and QLQ-BN20 quality of life questionnaire)
- Surgery associated neurological deficits (National Institutes of Health Stroke Scale)
- Adverse events (classified according to the US National Cancer Institute common toxicity criteria version 4.0)
- Survival time (days)

Nature and extent of the burden and risks associated with participation, benefit and group relatedness:

The standard treatment of glioblastoma patients consist of combined radiation and chemotherapy for a period of 6 weeks after surgery, followed by monthly cycles of chemotherapy alone, mostly during 6 months. Maximal and safe resection is the main goal of glioblastoma surgery and is currently still a great challenge. Finding an adjuvant neurosurgical tool to improve the extent of glioblastoma resection is of high importance. Ultrasound guided resection is a very low-risk alternative to a standard tumor resection procedure, which possibly shortly extends the time of surgery. In this study, patients will be randomised in two treatment arms: ultrasound guided or non-ultrasound guided glioblastoma surgery. During a follow up time of 6 months, patients will be called up 3 times to fill in 2 questionnaires to compare quality of life and neurological functioning. Both treatment arms will follow the standard treatment protocol for glioblastoma and no additional interventions will be done.

Introduction and rationale

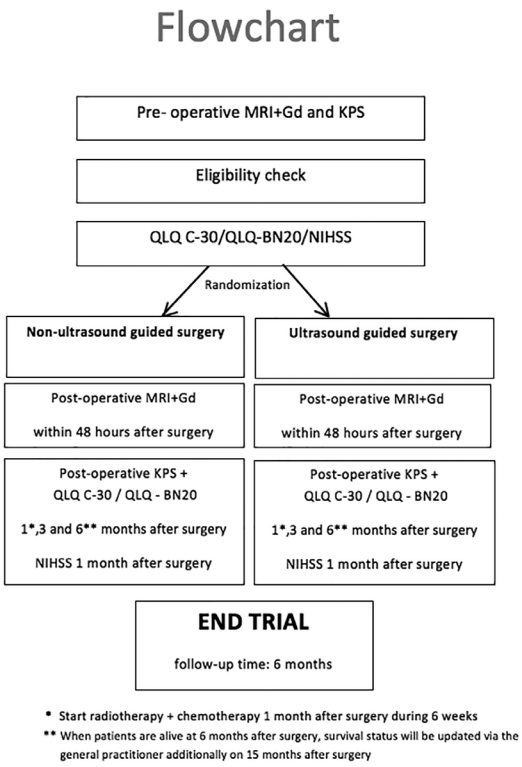
median survival of high grade glioma is only 15 months after surgery, radio- and chemotherapy[1]. Prognosis of patients with HGG is independently associated with larger surgical resections of the tumour. However, larger resections also have the risk of damaging normal brain and could therefore have detrimental effect on quality of life of these patients. Achieving GTR without causing new neurological deficits is therefore still a great challenge in glioma surgery[13-17].

The term gross total resection (GTR) is used in the literature to indicate no residual tumour on post-operative MRI scans. GTR in newly diagnosed, untreated HGG patients varies in the literature from 33% to 85% in retrospective studies. GTR of HGG patients operated in Erasmus MC is around 30% and comparable to the only known published randomized controlled study, investigating the use of fluoro-guided resection[16]. The relatively low GTR's are caused by the technical difficulty to identify interface during resection between tumour and white matter of the brain. The neurosurgeon is therefore not able to distinguish tumour from white matter clearly. New intraoperative high resolution imaging is needed to overcome this problem.

With the use of navigation equipment, pre-operative MRI scans are used to help neurosurgeons navigate to the destination of the glioma in the brain during surgery. However, due to brain and tumour shift during operation, these images do not reflect the real-time situation in the brain during surgery. Intraoperative acquired, real-time images are therefore needed to correct for brain and tumour shift to optimize tumour resection. Nowadays, intraoperative MRI (iMRI) is an increasingly used as a tool to acquire real-time images and to improve extent of glioma resection during surgery[15]. However, installing this technique in the operating room is highly expensive (between 3–5 million

euros), has high maintenance costs, is extremely time consuming during operations (1-2 hours), and will not be available for most of the neurosurgical centers around the world[15]. Intraoperative navigated high resolution ultrasound (US) could be an alternative.

Recently, a new intraoperative navigated high resolution US (developed by Brainlab and BK-medical) has been shown to be a promising **cost-effective** tool to acquire real-time intraoperative images to localize and to resect gliomas[18-22]. Intraoperative US guidance costs a fraction of an MRI, has almost no maintenance costs, is much less time consuming and can be used to acquire real time images during surgery[23-29]. With the use of (older) US, GTR percentages between 63%[11] and 94%[12] could be attained, but most of these published studies are biased by selection, where superficial small tumours have better resections as compared to more difficult to operate deeply seated larger tumours. The new intraoperative navigated high resolution US (developed by Brainlab and BK-medical) is a significant improvement as and has the potential to be incorporated standardly during resection of HGG in contrast to older (non-navigation fused) US devices. The purpose of this study is to investigate the effectivity of the newly acquired intraoperative navigated US in achieving GTR in patients with HGG and to measure influence on quality of life. Our primary goal is to investigate whether the use of IOUS additional to neuronavigation improves after the extent of high grade glioma resection compared with the use of neuronavigation only. Our secondary goal is to investigate whether ultrasound guided tumor resection improves neurological outcome, quality of life and survival time, when compared with non-ultrasound guided tumor resection



Objectives

Primary Objective

To investigate whether ultrasound guided tumor resection succeeds gross total resection significantly more frequently, when compared with the conventional non-ultrasound guided tumor resection.

Secondary Objective(s)

To investigate whether ultrasound guided tumor resection improves the extent of resection , quality of life and survival time, without causing new neurological deficits more frequently, when compared with the conventional non-ultrasound guided tumor resection.

Study design

Randomized Controlled Trial, Single Blinded

All newly diagnosed, untreated, contrast enhancing presumed high-grade glioma patients (18 years or older) will be randomized in two groups. In one group glioma surgery will be performed by using neuronavigation only (conventional treatment). In the other group, glioma surgery will be performed with the use of intraoperative ultrasound in addition to neuronavigation (ultrasound guided).

Primary outcome

The extent of resection will presented as a dichotomous outcome: gross-total resection or sub-total resection.

Gross-total resection: No residual contrast enhancement on post-operative MRI scans (within 48 hours); 100% of all contrast enhancing tumor has been resected when compared to initial enhancing tumor on pre-operative MRI scans. Sub-total resection: Residual contrast enhancement on post-operative MRI scans (within 48 hours); <100% of all contrast enhancing tumor has been resected when compared to initial enhancing tumor on pre-operative MRI scans.

Secondary outcome

The extent of resection (%) is a secondary outcome measurement defined as the residual tumor volumes on post-operative MRI studies compared to the operative tumor volume.

Patients will be followed for 6 months and the neurological outcome (KPS), Quality of Life (EORTC QLQ-C30,QLQ-BN20), surgery associated neurological deficits (measured 1 month after date of surgery), adverse events and time of survival (days) will be assessed as a secondary outcome measurement.

Schedule

	Within 1 week prior to treatment	Within 2 days after surgery	1 month after surgery2	3 months after surgery	6 months after surgery
Written informed consent	x				
Eligibility check	x				
MRI with contrast	x	x			
KPS	x		x	x	x
NIHSS	x		x		
QLQ-C30	x		x	x	x
QLQ-BN20	x		x	x	x
Survival			x	x	x
AE	x		x[8]		

KPS Karnofsky Performance Status, NIHSS National Institutes of Health Stroke Scale, QLQ-C30 and QLQ-BN20 EORTC quality of life questionnaires, AE adverse events.

Study population

population (base)

Subjects include contrast enhancing presumed high-grade glioma patients of 18 years or older, which are already selected for surgery with the goal of maximal tumor resection. Between January 2013 and October 2014, 148 patients with high grade glioma were operated at the Erasmus Medical Center hospital. Of all 148 patients, 49 patients matched the in- and exclusion criteria as written below. Based on this data, we expect to include all patients needed for this study (n=50) approximately within 18 months of time.

Inclusion criteria

In order to be eligible to participate in this study, a subject must meet all of the following criteria:

- Individuals of 18 years or older
- Newly diagnosed, untreated, contrast enhancing presumed high-grade glioma
- KPS ≥ 60
- Preoperative intention to perform gross-total resection of the enhancing tumor
- Written informed consent conform ICH-GCP

Exclusion criteria

- A potential subject who meets any of the following criteria will be excluded from participation in this study:
- Tumours crossing the midline basal ganglia, cerebellum, or brain stem prohibiting gross total resection
- Multifocal contrast enhancing lesions

- Pre-existing neurological deficit (e.g. aphasia, hemiparesis) due to neurological diseases (e.g. stroke)
- Inability to give consent because of dysphasia or language barrier

Sample size calculation

References	GTR* Non- US	GTR US**	Mean***	Definition GTR	N	Lesion volume	Pathology
Liang 2013[9]		90%	84%	100%	80	unclear	Glioma
Serra 2012[10]		93%		100%	14	Contrast enhancing	HGG****
Solheim 2010[11]		63%		<0.175 cm^3	57	Contrast enhancing	HGG
Tian 2009[12]		94%		unclear	58	unclear	HGG
Benveniste 2003[13]	85%		50%	100%	54	Contrast enhancing	HGG
Mc Girt 2009[14]	55%			100%	306	Contrast enhancing	HGG
Senft 2011[15]	68%			<0.175 cm^3	25	Contrast enhancing	HGG
Stummer 2006[16]	36%			<0.175 cm^3	131	Contrast enhancing	HGG
Wu 2007[17]	33%			100%	42	Contrast enhancing	HGG
EMC experience	25%			100%	49	Contrast enhancing	HGG

* Gross-total resection
** Ultrasound guided resection- study
*** (((GTR_1%) x (N_1)) + ((GTR_2%) x (N_2)) + ((GTR_x%) x (N_x))) / (N totaal)
**** High grade glioma

Based on the experience in the Erasmus MC and Stummer[16], we expect that 35% of the patients in the standard treatment arm will reach total resection. A target number of 23 patients in each treatment arm will give a power of 80% to detect an increase of 36% to 80% of the proportion of patients with a total resection at a significance level of 5%. To account for the possibility of drop-out and missing data of subjects, we will increase the sample size to 25 subjects per treatment arm, i.e. a total of 50 subjects.

Treatment of subjects

The standard treatment for glioblastoma is surgery or biopsy, followed by combined radiotherapy (RT) and chemotherapy. Surgery is performed to achieve maximal and safe tumor resection. Fractionated RT with a total of 60Gy is usually combined with Temozolomide (TMZ) (75mg/m2) daily.

During RT TMZ is taken orally daily within 2 h prior to RT starting on the first day of RT until 42 day since start RT. TMZ is taken fasting in the morning during the week and prior to breakfast at the weekends, when no RT will be delivered. This schedule is as registered. When TMZ is continued in adjuvant setting, it will be administered as 150 mg/m2 in cycle 1, and 200 mg/m2 in cycles 2-6, during 1-5 of each 4 week cycle, starting 4 weeks after

completion of RT. Standard adjuvant treatment with TMZ will consist of 6 cycles (and up to a maximum of 12 cycles).

Investigational product/treatment

Not applicable. All of the participants in this study will receive standard therapy after surgery, despite of in which treatment arm the subjects will be randomised for. This standard therapy is no subject of this study.

Based on the randomisation, the tumor will be removed surgically with ultrasound guidance (intervention) or without ultrasound guidance (standard/comparison).

During ultrasound guided resection, the neurosurgeon will use ultrasound for acquiring updated images during tumor removal. Before finishing resection and closing the dura, the surgeon will check for residual tumor. If there is no residual tumor resection will be finished and the dura will be closed. When residual tumor is seen, the surgeon will continue resection till all residual tumor has been resected.

Use of co-intervention NOT APPLICABLE
Escape medication NOT APPLICABLE

Investigational product

name and description of investigational product(s) NOT APPLICABLE
Summary of findings from non-clinical studies NOT APPLICABLE
Summary of findings from clinical studies NOT APPLICABLE
Summary of known and potential risks and benefits
Description and justification of route of administration and dosage NOT APPLICABLE
Dosages, dosage modifications and method of administration NOT APPLICABLE
Preparation and labelling of Investigational Medicinal Product NOT APPLICABLE
Drug accountability NOT APPLICABLE

Non-investigational product

Ultrasound is an intraoperative adjunct which is already registered and used by neurosurgeons in daily practice when necessary during tumor resection. Although this tool is already registered, we provided additional information about this product in this chapter. Patients will receive Temozolomide and radiotherapy, regardless of randomisation group, as a standard treatment protocol for GBM . We consider Temozolomide and radiotherapy as a background treatment (NIMP) following the definition of the EU *"THE RULES GOVERNING MEDICINAL PRODUCTS IN THE EUROPEAN UNION VOLUME 10 - GUIDANCE DOCUMENTS APPLYING TO CLINICAL TRIALS GUIDANCE ON INVESTIGATIONAL MEDICINAL PRODUCTS (IMPS) AND 'NON INVESTIGATIONAL MEDICINAL PRODUCTS' (NIMPS) (REV. 1,MARCH 2011)". S*ince Temozolomide/radiotherapy is already registered as a standard

therapy for GBM and we do not expect any kind of action of ultrasound on Temozolomide/ radiotherapy, we did not mention these background treatments in this chapter.

Name and description of non-investigational product(s): **BK Medical Flex Focus 800**

Specifications
Image Modes: B, M, Color Doppler, PW, Doppler, Tissue Harmonic
Features and Options: DVD RW, DICOM, BK Power Pack
Display: 19" LCD monitor
Dimensions: System height: 13050-1602 mm, Keyboard 745-1055mm, Body width: 350 mm, Depth: 610 mm
Weight: 49 kg, imaging unit only: 7 kg

BK Medical - Craniotomy Transducer 8862

Specifications
Frequency Range 10-3.8 MHz
Focal Range 5-68 mm
Contact surface 29x10 mm
Sector angle 66 degrees
Physical data 138 x 25 x 12 mm
Weight: 50 g
Summary of findings from non-clinical studies NOT APPLICABLE

Summary of findings from clinical studies
Intraoperative 3-D ultrasound seems to provide a safe, time- and cost-effective way to acquire real time images during surgery. The new navigated high resolution 3D US is at least as reliable as navigated 3D MR to delineate gliomas and metastases[18-22]. Studies have also suggested that US guidance is a successful technique to achieve gross total resection (GTR). GTR during high grade glioma resection without US guidance is achieved between 25% and 85% [13-17] (mean 50%) of cases published in several studies.

To the best of our knowledge, no randomized controlled trial has been done to investigate the value of US guidance during high grade glioma surgery. However, several studies have presented a higher GTR percentage when US was used during surgery. GTR was achieved in 6%-94%[9-12] of high grade glioma patients with the use of intraoperative ultrasound in several studies. Because of the different definitions of GTR in these studies (100% resection / <0.175 cm^3 residual tumor), different pathological (glioblastoma, (oligo) asterocytoma grade III-IV) and radiological (enhancing/non enhancing) characteristics of patients, different inclusion criteria in these studies, the results should be compared and interpreted with caution. Randomized controlled trials are highly needed to state the real

impact of using US guidance during high grade glioma resection, not only in terms of improving EOR, but also to state the real impact on neurological outcome, quality of life and survival time. According to the results of a meta-analysis of all major clinical publications since 1990 on the topic of the effect of improving EOR on neurological outcome, extensive surgical resection is associated with a longer life expectancy for patients with both LGG and high-grade gliomas.[30] Stummer et al.[3] showed that an improvement of EOR by using 5 ALA also improved survival time of patients with glioblastoma; patients undergoing complete tumor resection did significantly better than patients with residual tumor (50% survival rate at 57.8 weeks vs. 33.8 weeks, log rank test p=0.003).

Summary of known and potential risks and benefits

Ultrasound is a non-ionizing radiation, therefore it has an excellent safety record without any known risks. Based on the guideline by the NFU (Dutch Federation of University Medical Centers) about quality insurance in human research ("Kwaliteitsborging van mensgebonden onderzoek") we qualify the risk of this study as 'low' (small chance of serious damage).

Description and justification of route of administration and dosage NOT APPLICABLE
Dosages, dosage modifications and method of administration NOT APPLICABLE
Preparation and labelling of Non Investigational Medicinal Product NOT APPLICABLE
Drug accountability NOT APPLICABLE

Methods

Main study parameter/endpoint

- Extent of resection (gross-total resection or sub-total resection)

Gross-total resection: No residual contrast enhancement on post-operative MRI scans (within 48 hours); 100% of all enhancing tumor has been resected when compared to initial enhancing tumor on pre-operative MRI scans.
Sub-total resection: Residual contrast enhancement on post-operative MRI scans (within 48 hours); <100% of all enhancing tumor has been resected when compared to initial enhancing tumor on pre-operative MRI scans.

Secondary study parameters/endpoints

- Extent of resection (%)

Initial and residual tumor volume (cm^3) of all enhancing tissue on respectively pre and post-operative MRI scans will be volumetrically assessed.

The extent of resection (%) will be calculated with the formula: (initial tumor volume-residual tumor volume)/initial tumor volume x 100 [31].

- Neurological status (KPS) (Pre-operative and post-operative on 1, 3, 6 months) [32]
- Quality of Life (QLQ-C30 and QLQ-BN 20 questionnaire) (Pre and post-operative on 1, 3, 6 months) [33]
- Surgery associated neurological deficits (National Institutes of Health Stroke Scale, NIHSS) (Pre-operative and 1 month post-operative)[16,34]
- Adverse Events (classified according to the US National Cancer Institute common toxicity criteria version 4.0)
- Time of survival (days)

Other study parameters NOT APPLICABLE

Randomisation, blinding and treatment allocation

After eligibility check, participants will be randomized (1:1) to receive glioma resection with ultrasound guidance or without ultrasound guidance, using a web based randomisation program.

Single blinding

The radiologist who will assess pre- and post-operative tumor volumes as a primary outcome measurement will be blinded for patient's randomization result and surgical procedures applied. Data collection of the primary endpoint will be performed according to the study specific datamanagement plan.

Study procedures

All subjects will undergo the following procedures during this study:

- Pre operative MRI scan (part of the standard medical treatment)
- Post operative (within 48 h) MRI scan (part of the standard medical treatment)

- Pre operative KPS scoring
- Pre operative QLQ-C30 and QLQ-BN20 questionnaire
- Pre operative NIHSS

- 1 month post-operative KPS scoring
- 1 month post-operative QLQ-C30 and QLQ-BN20 questionnaire
- 1 month post-operative NIHSS

- 3 months post-operative KPS scoring
- 3 months post-operative QLQ-C30 and QLQ-BN20 questionnaire

- 6 months post-operative KPS scoring
- 6 months post-operative QLQ-C30 and QLQ-BN20 questionnaire

- When patients are alive at 6 months after surgery, survival status will be updated via the general practitioner additionally on 15 months after surgery.

Patients will be randomized into two treatment arms:
- Intraoperative neuronavigation guided glioma resection (conventional treatment) or
- Ultrasound integrated neuronavigation guided glioma resection (intervention)

The neurosurgeon will use the transducer to acquire ultrasound images during resection and before closing the dura to see if there is residual tumor left. The surgeon will continue with the resection till the absence of residual tumor on ultrasound images.

Withdrawal of individual subjects
Subjects can leave the study at any time for any reason if they wish to do so without any consequences. The investigator can decide to withdraw a subject from the study for urgent medical reasons.

Specific criteria for withdrawal NOT APPLICABLE
Replacement of individual subjects after withdrawal
All included patients will be analysed following an intention to treat analysis
Follow-up of subjects withdrawn from treatment
After withdrawal or end of the study the neurosurgeon (or if preferred by patient the neuro-oncologist) stays in contact with patient and takes care of further follow up conform standard practice if applicable
Premature termination of the study **NOT APPLICABLE**

Safety reporting

section 10 WMO event
In accordance to section 10, subsection 1, of the WMO, the investigator will inform the subjects and the reviewing accredited METC if anything occurs, on the basis of which it appears that the disadvantages of participation may be significantly greater than was foreseen in the research proposal. The study will be suspended pending further review by the accredited METC, except insofar as suspension would jeopardise the subjects' health. The investigator will take care that all subjects are kept informed.

AEs, SAEs and SUSARs

Adverse events (AEs)

An Adverse Event (AE) is any unfavorable and unintended sign or symptom, whether or not considered related to ultrasound guidance during tumor resection. Adverse events will be collected on the study specific CRF, which should be completed at baseline and at 1 month after surgery, i.e. before start of standard chemoradiation.

Serious adverse events (SAEs)

- A serious adverse event is any untoward medical occurrence or effect that at any dose:
- results in death;
- is life threatening (at the time of the event);
- requires hospitalisation or prolongation of existing inpatients' hospitalisation;
- results in persistent or significant disability or incapacity;
- is a congenital anomaly or birth defect;
- Any other important medical event that may not result in death, be life threatening, or require hospitalization, may be considered a serious adverse experience when, based upon appropriate medical judgement, the event may jeopardize the subject or may require an intervention to prevent one of the outcomes listed above.

The term 'life-threatening' in the definition of 'serious' refers to an event in which the subject was at risk of death at the time of the event. It does not refer to an event, which hypothetically might have caused death, if it were more severe. In general, 'hospitalization' signifies that the subject has been detained (usually involving at least an overnight stay) at the hospital or emergency ward for observation and/or treatment that would not have been appropriate in the physician's office or out-patient setting. Complications that occur during hospitalization are AEs, if a complication prolongs hospitalization or fulfils any other serious criteria, the event is serious. Hospitalization for elective treatment of a pre-existing condition that did not worsen fromm baseline does not meet the SAE criteria. Situations in which an untoward medical occurrence did not occur (e.g. social and/or convenience admission to a hospital) do not meet the SAE criteria and should therefore not be reported as SAEs. When in doubt as to whether "hospitalization" occurred or was necessary, the AE should be considered serious.

The term 'disability' means a substantial disruption of a person's ability to conduct normal life functions. This definition is not intended to include experiences of relatively minor medical significance such as uncomplicated headache, nausea, vomiting, diarrhea, influenza, and accidental trauma (e.g. sprained ankle) which may interfere or prevent everyday life functions but do not constitute a substantial disruption. The sponsor will report the SAEs occurring between randomisation date and 30 days after surgery through

the web portal *ToetsingOnline* to the accredited METC that approved the protocol, within 15 days after the sponsor has first knowledge of the serious adverse events, regardless the causal relationship to ultrasound guided surgery. SAEs that result in death or are life threatening should be reported expedited. The expedited reporting will occur not later than 7 days after the responsible investigator has first knowledge of the adverse event. This is for a preliminary report with another 8 days for completion of the report.

Suspected unexpected serious adverse reactions (SUSARs) NOT APPLICABLE
Annual safety report NOT APPLICABLE

Follow-up of adverse events
All AEs will be followed until they have abated, or until a stable situation has been reached. Depending on the event, follow up may require additional tests or medical procedures as indicated, and/or referral to the general physician or a medical specialist.
SAEs need to be reported till end of study within the Netherlands, as defined in the protocol

[Data Safety Monitoring Board (DSMB) / Safety Committee]
During this study no DSMB board will be installed.

Statistical analysis

primary study parameter(s)
The difference in the percentage of patients with GTR between trial arms will be analysed using continuity-adjusted chi squared test. The data of all included (eligible) patients will be analysed according to the intention-to-treat principle.

Secondary study parameter(s)
The extent of resection (%) as a continuous variable will be compared between trial arms using a Mann-Whitney test.

Quality of Life Assessment
The European organisation for research and treatment of cancer (EORC) developed the QLQ-C30 questionnaire for cancer patients and the disease specific QLQ-BN20, specifically developed and validated for patients with brain tumor[33]. Both tools have been tested and validated in clinical trials. The 50 questions in both questionnaires together take 10 minutes to complete. The EORTC QLQ-C30 measures functioning scales – physical, role, emotional, cognitive and social; three symptom scales – fatigue, nausea/ vomiting and pain; six single item scales – dyspnoea, insomnia, appetite loss, constipation, diarrhoea and financial impact and the overall HRQOL-scale.

The EORC QLQ-BN20 is designed for patients undergoing chemotherapy or radiotherapy and includes 20 items assessing visual disorders, motor dysfunction, communication deficit, various disease symptoms (e.g. headaches and seizures), treatment toxicities (e.g. hair loss), and future uncertainty.

Both the quality of life questionnaires EORTC QLQ-C30 and the EORTC QLQ-BN20 will be scaled, scored and transformed to a linear scale (0-100). Differences ≥10 points are classified as clinically meaningful changes in a HRQL parameter. Changes > 20 points are classed as large effects. The questionnaires will be filled out by the patient before surgery and on 1, 3 and 6 months after surgery.

Neurological outcome (KPS), surgery associated neurological deficits (NIHSS) and QoL (QLQ-C30 and QLQ-BN20) will be presented as continuous variables, with means and standard deviations (if normally distributed) or medians and interquartile ranges (if not normally distributed). These data will be analysed using linear mixed models, with time since baseline, study arm, the interaction effect between time since baseline and study arm, and other relevant clinical factors as predictor variables. In the model, we will impose that the mean scores of KPS, NIHSS and QoL are not significantly different at randomization. If KPS and/or NIHSS and/or QoL are not normally distributed, then an appropriate transformation to normality may be applied before including these variables as the dependent variable in the linear mixed models. Overall survival will be analysed with Cox regression analysis. Also a Kaplan-Meier curve will be presented (by treatment arm). The analysis of treatment toxicities will primarily done by tabulation of the incidence of adverse events CTCAE grade 2 or more by treatment arm.

Other study parameters NOT APPLICABLE
Interim analysis NOT APPLICABLE

Ethical considerations

Regulation statement

This study will be conducted according to the principles of the Declaration of Helsinki (19 October 2013) and in accordance with the Medical Research Involving Human Subjects Act (WMO) and other guidelines, regulations and Acts.

Recruitment and consent

The neurosurgeon (who will perform tumor resection) will inform the subject about the study. The investigator will give the patient a patient information letter and informed consent letter and will ask their consent. Patients are able to consider their decision for at least 1 day.

Objection by minors or incapacitated subjects *NOT APPLICABLE*

Benefits and risks assessment, group relatedness
As maximal and safe glioblastoma resection is still a challenge, it is of high importance to find an adjuvant intra-operative tool to safely maximize the extent of resection of glioblastoma resection.

Administrative aspects, monitoring and publication

handling and storage of data and documents

Patient confidentiality
Each patient is assigned a unique patient study number at enrolment. In trial documents the patient's identity is coded by patient study number as assigned at enrolment.

The local investigator will keep a subject enrolment and identification log that contains the key to the code, i.e. a record of the personal identification data linked to each patient study number. This record is filed at the investigational site and should only be accessed by the investigator and the supporting site staff, and by representatives of the Sponsor or a regulatory agency for the purpose of monitoring visits or audits and inspections.

Case Report Forms (CRF)
Data will be collected in an electronic CRF to document eligibility, safety and efficacy parameters, compliance to treatment schedules and parameters necessary to evaluate the study endpoints. Data to be collected on the CRF are derived from the protocol.

Radiological images pre- and post-surgery will be stored on a secure server location that will be used for the central radiologists to assess the primary endpoints. The primary outcome measures will be recorded by the independent radiologist on a paper CRF that will not include randomization result but only the patient specific subject number. The completed CRF's will be entered in the database by the Clinical Trial Center.

Filling of essential documents
Essential Documents are those documents that permit evaluation of the conduct of a trial and the quality of the data produced. The essential documents may be subject to, and should be available for, audit by the Sponsor's auditor and inspection by the regulatory authority(ies)
The investigator should file all essential documents relevant to the conduct of the trial on site. The Sponsor will file all essential documents relevant to the overall conduct of the trial. Essential documents should be filed in such a manner that they are protected from accidental loss and can be easily retrieved for review.

Record retention
Essential documents should be retained for 15 years after the end of the trial (i.e. from date of last patient visit for this trial). They should be destroyed after this time. Source documents (i.e. medical records) of patients should be retained for at least 15 years after the end of the trial. Record retention and destruction after this time is subject to the site's guidelines regarding medical records.

Monitoring
On behalf of the Sponsor the Clinical Trial Center will perform on-site monitoring visits to verify that the rights and well-being of patients are protected, the reported trial data are accurate, complete, and verifiable from source documents and the conduct of the trial is in compliance with the currently approved protocol/amendment(s), with GCP, and with the applicable regulatory requirement(s). Monitoring visits will take place according to the study specific monitoring plan.

On-site monitoring includes checking informed consent procedures, timely reporting of SAEs, and verification of completeness of the Investigator Site File, conform the study specific monitoring plan. Minor and major findings of the monitor will be discussed with the local investigator, and documented in a standard monitoring report that will be provided to the Sponsor. The Sponsor may decide to increase the monitoring frequency or intensity if the results of monitoring require this to ensure patient safety and/or data quality.

Direct access to source documentation (medical records) must be allowed for the purpose of verifying that the data recorded in the CRF are consistent with the original source data.

The Clinical Trial Center will perform central monitoring on collected data, including checks on completeness of the data, data inconsistencies, timely reporting of SAEs etcetera, on a regular base. Local investigators will remain responsible for obtaining essential documents that needs to be filed in the Investigator Site File. In case major violations are found during the random source data verification, such as failure to report SAE's, the Sponsor may decide that additional monitoring is indicated.

Quality assurance
Steps to be taken to ensure the accuracy and reliability of data include the selection of qualified investigators and appropriate study centers, review of protocol procedures with the investigator before the study, and site visits by the Sponsor.
Data collected on the CRF will be verified for accuracy. If necessary, queries will be sent to the investigational site on a regular base to clarify the data on the CRF. The investigator or his/ her delegate should answer data queries within the specified time line.

Audits and inspections
The investigator will permit auditors to carry out site visits to audit the compliance with regulatory guidelines. These audits will require access to all study records, including source documents, for inspection and comparison with the CRFs. Patient privacy must, however, be respected. Similar auditing procedures may be conducted by agents of any regulatory body reviewing the results of this study. The investigator should immediately notify the Sponsor if they have been contacted by a regulatory agency concerning an upcoming inspection.

Amendments
Amendments are changes made to the research after a favourable opinion by the accredited METC has been given. All amendments will be notified to the METC that gave a favourable opinion.

Annual progress report
The sponsor/investigator will submit a summary of the progress of the trial to the accredited METC once a year. Information will be provided on the date of inclusion of the first subject, numbers of subjects included and numbers of subjects that have completed the trial, serious adverse events/ serious adverse reactions, other problems, and amendments.

End of study report
The investigator will notify the accredited METC of the end of the study within a period of 90 days. The end of the study is defined as the last patient's last visit. In case the study is ended prematurely, the investigator will notify the accredited METC within 15 days, including the reasons for the premature termination. Within one year after the end of the study, the investigator/sponsor will submit a final study report with the results of the study, including any publications/abstracts of the study, to the accredited METC.

Public disclosure and publication policy
Publications resulting from this study will be submitted to peer-reviewed journals. The principle investigators and study coordinators will prepare the manuscript together with those who substantially contributed to the study. Registration of the clinical trial will be done in a public trial registry before the first patient is recruited.

Structured risk analysis

a. Level of knowledge about mechanism of action Ultrasound, is high-frequency sound. It is generated by a transducer that converts electrical signals into ultrasound waves and picks up the reflected signals converting them back into electrical signals. These

signals are then seen on the screen. The detailed mechanism of action of ultrasound is well described in earlier published literature[35].

b. Previous exposure of human beings with the test product(s) and/or products with a similar biological mechanism *NOT APPLICABLE*
c. Can the primary or secondary mechanism be induced in animals and/or in *ex-vivo* human cell material? *NOT APPLICABLE*
d. Selectivity of the mechanism to target tissue in animals and/or human beings Attenuation describes the loss of energy, expressed as change in intensity, as sound waves travel through a medium. Ultrasound is reflected at the boundaries between different materials. In neurosurgery, this means that if the consistency between two different tissues (for example, normal brain and tumor) is significantly different, the lesion will be clearly visible. However, if the tumor tissue has a similar consistency to the normal brain (as in some LGGs) it is more difficult to distinguish the difference[35].
e. Analysis of potential effect The brain is a visco-elastic medium, and the ultrasound waves are not only propagated through the tissue, but are also transformed into heat and absorbed. This resultant loss of energy produces a darker far field image without compensation[35].
f. Pharmacokinetic considerations *NOT APPLICABLE*
g. Study population Subjects include stable high-grade glioma patients between the age of 18-75 years, which are already selected and involved in surgically glioma resection care.
h. Interaction with other products *NOT APPLICABLE*
i. Predictability of effect NOT APPLICABLE
j. Can effects be managed? NOT APPLICABLE

Synthesis

Ultrasound is an intra-operative adjunct which is already registered and used by neurosurgeons in daily practice when necessary during tumor resection. Adding ultrasound to the standard surgery procedure gives minimal risks for patients. The neurosurgeons involved in this study have had sufficient exposure to patients with high grade glioma. Based on the guideline by the NFU (Dutch Federation of University Medical Centers) about quality insurance in human research ("Kwaliteitsborging van mensgebonden onderzoek") and Erasmus MC Policy on monitoring, we qualify the risk of this study as 'low' (small chance of serious damage).

References

1. 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. J Neurosurg 95:190–198, 2001
2. Tsitlakidis A, Foroglou N, Venetis CA, Patsalas I, Hatzisotiriou A, Selviaridis P. Biopsy versus resection in the management of malignant gliomas: a systematic review and meta-analysis. J Neurosurg. 2010 May;112(5):1020-32.
3. Stummer W, Reulen HJ, Meinel T, Pichlmeier U, Schumacher W, Tonn JC et al. ALA-Glioma Study Group. Extent of resection and survival in glioblastoma multiforme: identification of and adjustment for bias. Neurosurgery. 2008 Mar;62(3):564-76; discussion 564-76.
4. Chaichana KL, Jusue-Torres I, Navarro-Ramirez R, Raza SM, Pascual-Gallego M, Ibrahim A et al. Establishing percent resection and residual volume thresholds affecting survival and recurrence for patients with newly diagnosed intracranial glioblastoma. Neuro Oncol. 2014 Jan;16(1):113-22.
5. McGirt MJ, Chaichana KL, Gathinji M, Attenello FJ, Than K, Olivi A et al. Independent association of extent of resection with survival in patients with malignant brain astrocytoma. J Neurosurg. 2009 Jan;110(1):156-62.
6. Graus F, Bruna J, Pardo J, Escudero D, Vilas D, Barceló I et al. Patterns of care and outcome for patients with glioblastoma diagnosed during 2008-2010 in Spain. Neuro Oncol. 2013 Jun
7. Bauchet L, Zouaoui S, Darlix A, Menjot de Champfleur N, Ferreira E, Fabbro M et al.. Assessment and treatment relevance in elderly glioblastoma patients. Neuro Oncol. 2014 May 2.
8. Filippini G, Falcone C, Boiardi A, Broggi G, Bruzzone MG, Caldiroli D et al.; Prognostic factors for survival in 676 consecutive patients with newly diagnosed primary glioblastoma. Neuro Oncol. 2008 Feb;10(1):79-87.
9. Liang SQ, Kang JMl. Value of intraoperative ultrasound in precise glioma resection. Zhonghua Yi Xue Za Zhi. 2013; 93(33):2671-3.
10. Serra C, Stauffer A, Actor B, Burkhardt JK, Ulrich NH, Bernays RL et al. Intraoperative high frequency ultrasound in intracerebral high-grade tumors. Ultraschall Med. 2012;33(7):E306-12.
11. Solheim O, Selbekk T, Jakola AS, Unsgard G. Ultrasound-guided operations in unselected high-grade gliomas--overall results, impact of image quality and patient selection. Acta Neurochir (Wien). 2010 ;152(11):1873-86.
12. Tian YJ, Lin S, Liu HZ, Wang LS, He W, Zhang MZ, et al. Value of intra-operative ultrasound in detecting the boundaries of intra cranial gliomas. Zhonghua Yi Xue Za Zhi. 2009;89(19):1305-8.
13. Benveniste R, Germano IM. Evaluation of factors predicting accurate resection of highgrade gliomas by using frameless image-guided stereotactic guidance. Neurosurg Focus. 2003: 14 (2): e5
14. McGirt MJ, Mukherjee D, Chaichana KL, Than KD, Weingart JD, Quinones-Hinojosa A.. Association of surgically acquired motor and language deficits on overall survival after resection of glioblastoma multiforme. Neurosurgery. 2009 Sep;65(3):463-9.

15. Senft C, Bink A, Franz K, Vatter H, Gasser T, Seifert V. Intraoperative MRI guidance and extent of resection in glioma surgery: a randomised, controlled trial. Lancet Oncol. 2011 Oct;12(11):997-1003.
16. Stummer W, Pichlmeier U, Meinel T, Wiestler OD, Zanella F, Reulen HJ et al. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. Lancet. Oncol. 2006 May;7(5):392-401.
17. Wu JS, Zhou LF, Thang WJ, Mao Y, Hu J, Song YY et al. Clinical evaluation and follow-up outcome of diffusion tensor imaging-based functional neuronavigation: a prospective, controlled study in patients with gliomas involving pyramidal tracts. Neurosurgery. 2007;61(5):35–49
18. Coburger J, Scheuerle A, Thal DR, Engelke J, Hlavac M, Wirtz CR et al. Linear array ultrasound in low-grade glioma surgery: histology-based assessment of accuracy in comparison to conventional intraoperative ultrasound and intraoperative MRI. Acta Neurochir (Wien). 2015 Jan 6. [Epub ahead of print]
19. Tirakotai W, Miller D, Heinze S, Benes L, Bertalanffy H, Sure U. A novel platform for image-guided ultrasound. Neurosurgery. 2006 Apr;58(4):710-8; discussion 710-8.
20. Enchev Y, Bozinov O, Miller D, Tirakotai W, Heinze S, Benes L et al. Image-guided ultrasonography for recurrent cystic gliomas. Acta Neurochir (Wien). 2006 Oct;148(10):1053-63; discussion 1063.
21. Berntsen EM, Gulati S, Solheim O, Kvistad KA, Torp SH, Selbekk T et al. Functional magnetic resonance imaging and diffusion tensor tractography incorporated into an intraoperative 3-dimensional ultrasound-based neuronavigation system: impact on therapeutic strategies, extent of resection, and clinical outcome. Neurosurgery. 2010 Aug;67(2):251-64.
22. Lindner D, Trantakis C, Renner C, Arnold S, Schmitgen A, Schneider J et al. Application of intraoperative 3D ultrasound during navigated tumor resection. Minim Invasive Neurosurg. 2006 Aug;49(4):197-202.
23. Rasmussen IA Jr, Lindseth F, Rygh OM, Berntsen EM, Selbekk T, Xu J, et al. Functional neuronavigation combined with intra-operative 3D ultrasound: Initial experiences during surgical resections close to eloquent brain areas and future directions in automatic brain shift compensation of preoperative data. Acta Neurochir (Wien). 2007; 149: 365–378.
24. Unsgaard G, Ommedal S, Muller T, Gronningsaeter A, Nagelhus Hernes TA. Neuronavigation by Intraoperative Three-dimensional Ultrasound: Initial Experience during Brain Tumor Resection. Neurosurgery. 2002 Apr;50(4):804-12; discussion 812.
25. Gerganov VM, Samii A, Giordano M, Samii M, Fahlbusch R. Two-dimensional high-end ultrasound imaging compared to intraoperative MRI during resection of low-grade gliomas. Journal of Clinical Neuroscience. 2011; 18:669–673.
26. Erdogan N, Tucer B, Mavili E, Menku A, Kurtsoy A. Ultrasound guidance in intracranial tumor resection: correlation with postoperative magnetic resonance findings. Acta Radiol. 2005 Nov;46(7):743-9.
27. Unsgaard G, Selbekk T, Brostrup Muller T, Ommedal S, Torp SH, Myhr G,et al. Ability of navigated 3D ultrasound to delineate gliomas and metastases – comparison of image interpretations with histopathology. Acta Neurochir (Wien). 2005 Dec;147(12):1259-69

28. Tronnier V, Bonsanto MM, Staubert A, Knauth M, Kunze S, Wirtz CR. Comparison of intraoperative MR imaging and 3D-navigated ultrasonography in the detection and resection control of lesions. Neurosurg Focus. 2001;10:1–5.
29. Unsgaard G, Rygh OM, Selbekk T, Muller TB, Kolstad F, Lindseth F, et al. Intra-operative 3D ultrasound in neurosurgery. Acta Neurochir (Wien). 2006; 148: 235–253
30. Sanai N, Berger MS. Glioma extent of resection and its impact on patient outcome. Neurosurgery. 2008;62:753–64.
31. Smith JS, Chang EF, Lamborn KR, Chang SM, Prados MD, Cha S, et al., Role of extent of resection in the long-term outcome of low-grade hemispheric gliomas. J Clin Oncol. 2008; 26(8):1338-45
32. Cheng J, Zhang X, Liu B. Health-related quality of life in patients with high-grade glioma. Neuro-Oncology. 2009; 11, 41–50.
33. Taphoorn MJ, Claassens L, Aaronson NK, Coens C, Mauer M, Osoba D, et al., An international validation study of the EORTC brain cancer module (EORTC QLQ-BN20) for assessing health-releated quality of life and symptoms in brain cancer patients. Eur J Cancer. 2010; 46(6): p. 1033-40.
34. Khan RB, Gutin PH, Rai SN, Zhang L, Krol G, DeAngelis LM. Use of diffusion weighted magnetic resonance imaging in predicting early postoperative outcome of new neurological deficits after brain tumor resection. Neurosurgery. 2006 Jul;59(1):60-6; discussion 60-6.
35. Ivanov M, Wilkins S, Poeata I, Brodbelt A, et al. Intraoperative ultrasound in neurosurgery – a practical guide. British Journal of Neurosurgery.2010; 24(5): 510–517

Chapter 7

Clinical feasibility of a wearable mixed reality device in neurosurgery

Fatih Incekara, Marion Smits, Clemens Dirven, Arnaud J.P.E. Vincent

World Neurosurg. 2018 Oct;118:e422-e427

Abstract

Background

Neuronavigation systems are routinely used during neurosurgical procedures. Currently, new imaging technologies are emerging such as virtual, augmented and mixed reality. With mixed reality devices, the user can analyze and interact with the real environment using virtual objects. The aim of this prospective pilot study was to offer a proof of concept by testing the clinical feasibility and accuracy of a wearable mixed reality device (Hololens) for pre-operative neurosurgical planning.

Methods

In patients with an indication for brain tumor surgery, pre-operative planning of tumor localization with the Hololens was compared with standard neuronavigation in the operating room. Magnetic resonance imaging based 3D holograms of the patient's head and tumor were created and projected on the physical patient's head using the Hololens. The 2D projection of the tumor borders as perceived by the neurosurgeon on the skin of the patient's head was outlined both with the Hololens and neuronavigation. Accuracy of the Hololens localization was assessed using neuronavigation as the gold standard.

Results

Twenty-five patients were included in this study. Holograms were successfully created in all cases. In nine patients, tumor localization with the Hololens did not differ from the standard neuronavigation system and the overall median difference was 0.4 cm (IQR 0-0.8).

Conclusion

This prospective clinical study offers a proof of concept of the clinical feasibility of the Hololens for brain tumor surgery planning in the operating room, with quantitative outcome measures. Further development is needed to improve the accuracy of this wearable mixed reality device.

Introduction

In the early 1990s the first commercially available neuronavigation system was introduced.[1] Image guided navigation technologies have undergone several developments since then which have led to important improvements during both pre- and intra-operative neurosurgical procedures. Although limited evidence of an improvement of survival or quality of life of patients with neuronavigation guidance is available, neuronavigation is a highly practical and time efficient tool for pre- and intra-operative target localization.[2,3] Therefore, now almost 30 years later, the neuronavigation system has found its place in the operating room as a standard tool in cranial neurosurgery. Neuronavigation consists of several important applications, one of which is the pre-operative localization of the target lesion on the outer surface of a patient's head. However, this system has a considerable setup time due to a coordinate-based registration procedure. Additionally, the surgeon needs to translate these coordinates from a monitor to the patient's head for an accurate target localization.

Currently, we are seeing newly emerging imaging technologies which have a potential medical applicability by creating virtual objects in the real, physical world, and the ability of interaction between these objects and the real world. Virtual reality (VR), in which a real environment is replaced by a virtual one, and augmented reality (AR), in which virtual objects are overlaid on a real, physical environment, are such innovative technologies.[4,5] These new technologies are potentially of great benefit to neurosurgeons for complex surgical procedures. Meola et al. systematically reviewed AR studies within the field of neurosurgery and they concluded that AR has a potential benefit to improve current neuronavigation systems, but that prospective and clinical application studies are very limited.[6]

In contrast to VR and AR, mixed reality can be used to analyze and interact with the real environment by placing a virtual object over the physical environment. Our hypothesis is that mixed reality can be used for pre-operative planning and localization of the tumor by placing 3D hologram reconstructions of pre-operatively acquired brain MRI or CT scans over the physical operating field. To the best of our knowledge, a wearable mixed reality device has as yet not been tested within the field of (oncological) neurosurgery in a prospective clinical study. The aim of this study was to test the clinical applicability and accuracy of a wearable mixed reality device for pre-operative neurosurgical tumor localization and planning compared with standard neuronavigation.

Methods

Study participants

Patients with the age of 18 years and older, referred for neurosurgery of a newly diagnosed contrast enhancing intra- or extra-axial lesion as seen on a pre-operative T1 with contrast MRI scan, suspected for brain tumor, between December 2016 and December 2017 at the department of neurosurgery, Erasmus MC were eligible for this study. This study was approved by the IRB. Obtaining written informed consent was waived by the IRB, because participants were not subjected to any specific study procedures nor were they required to follow rules of behavior, and no additional risk was introduced by the study.

Hologram creation

Tumor volumes were semi-automatically segmented on high-resolution 3D contrast enhanced T1 weighted MRI images using a free online toolkit (ITK-SNAP; www.itksnap.org) and 3D objects (Figure 1) were reconstructed using free online software (Meshmixer; www.meshmixer.com). Then, 3D objects were sent online to the Hololens using a commercially available application (VertoStudio; www.vertostudio.com). The Microsoft Hololens is a commercially available, wearable computer integrated mixed reality device (www.microsoft.com/microsoft-hololens).

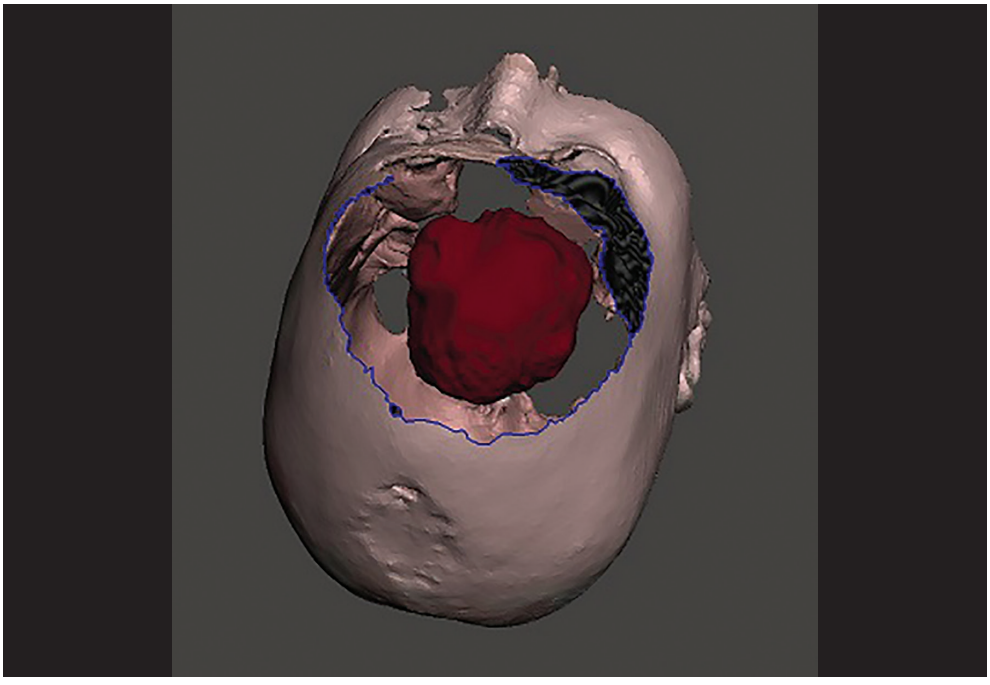


Figure 1. 3D reconstruction of the head of the patient with the edges of the virtual craniotomy (blue) and the tumor (red). This model is converted to a hologram and send to the mixed reality device.

Operative planning with the Hololens

From tumor segmentations, tumor volumes (cm^3) were measured using ITK-SNAP. Tumors were categorized as superficial (distance of 0 cm) or deep tumors (distance > 0 cm), as measured from the tumor border closest to brain surface to brain surface. Additionally, based on their localization, tumors where categorized by hemisphere (right vs. left) and lobe (frontal, parietal, temporal and occipital). Planning time for both the Hololens as the standard navigation system were measured from time to system setup to the final tumor outline on the patient's head. In the operating room, skin surface registration was performed manually by merging the 3D holograms of the patient's head and tumor with the physical head of the patient. Then, the 2D projection of the holographic tumor borders and the center of the tumor were marked on the skin of the patient's head as perceived by the neurosurgeon wearing the Hololens (Figure 2 and Figure 3). The center and outlines of the tumor were then marked again on the skin of the patient's head using the standard neuronavigation system (BrainLab, Feldkirchen, Germany). The maximal extension of the borders of the tumors were marked in 4 directions: anterior, posterior, medial and lateral. The exact center of the tumor projection on the skin was determined by the crossing of the anterior-posterior line and the medial-lateral line. The same was done with the standard neuronavigation system. As an outcome for accuracy in tumor localization, the maximum distance (cm) between the center of the tumor determined with the Hololens versus with neuronavigation (as the gold standard) was measured on the patient's skin. Eventually, skin incisions, craniotomy and surgery were performed based on tumor localization using the standard neuronavigation system. All measurements were performed by authors with prior Hololens experience (F.I and/or A.V). Cases were excluded from this study, if both of the authors could not attend surgery.

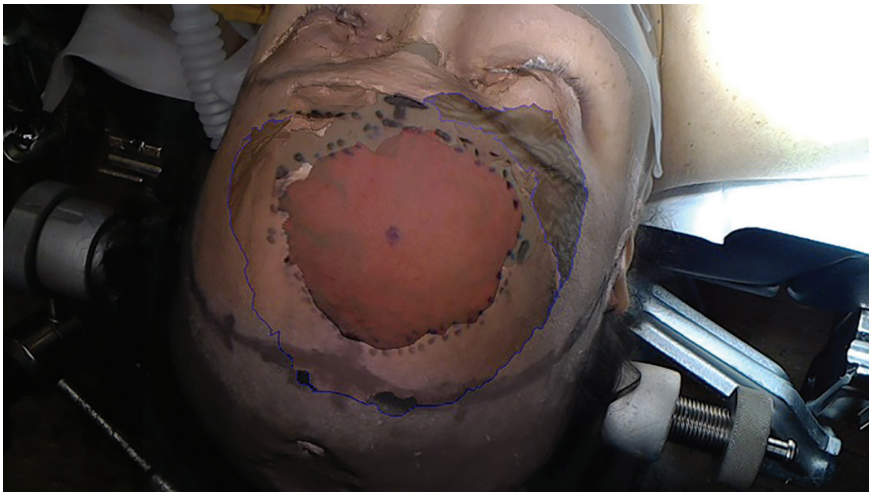


Figure 2. The head of the patient as seen through the Hololens. The 3D hologram is merged with the physical head of the patient. The tumor (red) is outlined first with the Hololens (light blue dots) and then with the navigation system (black dots). The two contours are in line with each other.

Figure 3. Neurosurgeon with the Hololens in the operating room.

Qualitative assessment

To obtain an impression of the alignment of the Hololens tumor projection in full, the holographic tumor borders of a meningioma were compared qualitatively with the true meningioma borders as seen on the brain surface, after opening the dura, in one patient. This assessment was also compared with the pre-operative tumor outlines on the skin.

Statistical Analysis

Statistical analyses were performed with the SPSS statistical software (SPSS version 21.0, SPSS Inc. Chicago, IL, USA). A one-sample Wilcoxon signed rank test was performed to test whether there was a significant deviation between Hololens and neuronavigation. Differences in deviation between groups of the various tumor locations between the Hololens and neuronavigation system were tested with a one-way ANOVA test. The difference in deviation between superficial and deep tumors and between the first and second half of the study, as well as the difference in pre-operative planning time between Hololens and standard neuronavigation were tested with a t-test. A p value of <0. 05 was considered to be statistically significant.

Results

Twenty-five patients were included in this study with a mean age of 57 years (range 22-80). Tumors had a median volume of 34.8 cm^3 (IQR 5.8 – 58.2) and 17 of 25 (68%) tumors reached the brain surface. Patient and tumor characteristics are further presented in Table 1. Holograms were successfully created in all of the patients. Mean pre-operative planning time with the Hololens (mean 5 min 20 sec, SD 1 min 20 sec) was longer, when compared with the standard neuronavigation system (mean 4 min 25 sec, SD 1 min 20 sec, $p < 0.001$). In nine patients (36%) tumor localization with the Hololens did not differ from that of the standard neuronavigation system. Overall, there was significant deviation between the Hololens and neuronavigation ($p < 0.0001$), with a median deviation of 0.4 cm (IQR 0-0.8). There was no statistically significant difference in deviation between tumor location in the left (median = 0.3 cm, IQR0-0.5) and right hemisphere (median = 0.5 cm, IQR 0-1.1, $p = 0.17$) or between the frontal, parietal, temporal and occipital lobes ($p = 0.74$). Additionally, no significantly different deviation was found between superficial (median = 0.4 cm, IQR 0-0.8) and deep tumors (median = 0.25 cm, IQR 0-0.8, p=0.65). There was a trend towards more accurately localizing tumors towards the end of the study (Figure 4), with a median deviation of 0.6 cm (IQR 0-1.1) during the first half (n=12) of the study and a median deviation of 0.3 cm (IQR 0-0.5) during the second half (n=13) of the study ($p = 0.07$). The qualitative assessment method showed that after opening the dura, the tumor as seen on the brain surface, was within the borders of the holographic tumor and also within the contour as marked on the patient's head using the Hololens.

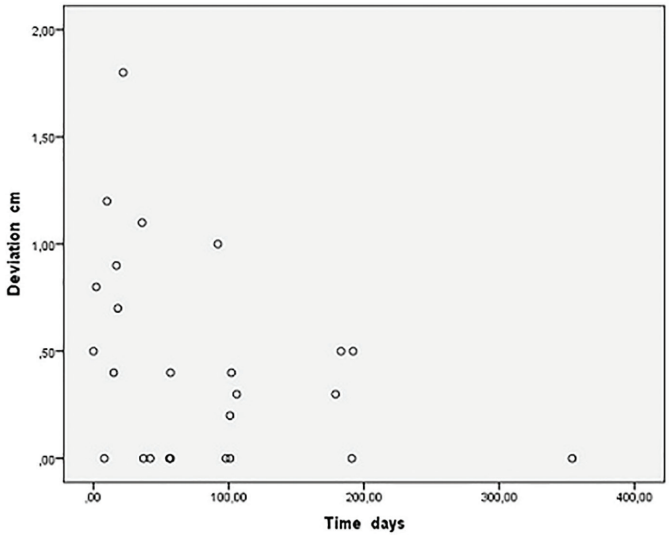


Figure 4. A scatterplot of deviation (cm) on the y-asis over time (days) on the x-asis, with t = 0 inclusion of first patient.

Table 1. Patient and tumor characteristics with deviation in tumor localization with Hololens, compared with neuronavigation system.

Age	Sex	Tumor localization	Tumor volume (cm3)	Tumor depth * (cm)	Pathology	Time pre-operative planning Hololens | navigation (min)	Deviation Hololens from navigation (cm)
71	Female	Right parietal	5.2	2.3	Glioblastoma	6 | 4	0.5
80	Male	Right temporal	9.4	0	Glioblastoma	5 | 3	0.8
38	Male	Left temporal	72.5	0	Meningioma	3 | 4	0
48	Female	Right frontal	29.1	0	Meningioma	7 | 5	1.2
67	Male	Left frontal	2.0	0	Glioblastoma	3 | 5	0.4
64	Male	Left temporal	94.4	1.2	Meningioma	5 | 4	0.9
54	Female	Right occipital	34.8	0	Metastasis	6 | 8	0.7
73	Male	Right frontal	1.8	0	Lymphoma	4 | 4	1.8
57	Male	Right temporal	6.4	24.0	Glioblastoma	5 | 4	1.1
53	Female	Left parietal	87.5	0	Glioblastoma	7 | 8	0
65	Male	Left temporal	35.1	1.3	Glioblastoma	7 | 4	0
70	Female	Left frontal	54.5	0	Metastasis	5 | 4	0
54	Male	Right frontal	45.5	0	Glioblastoma	5 | 3	0,4
73	Female	Right parietal	44.3	0	Glioblastoma	4 | 3	0
77	Female	Left frontal	32.9	0	Glioblastoma	3 | 3	1.0
63	Female	Left frontal	35.9	1.4	Glioblastoma	6 | 4	0
22	Female	Left frontal	36.2	0.5	Astrocytoma	5 | 5	0.2
63	Female	Right temporal	11.7	0.8	Meningioma	4 | 5	0
62	Female	Left frontal	61.9	0	Glioblastoma	5 | 5	0.4
28	Female	Left frontal	4.2	0.6	Astrocytoma	6 | 3	0.3
51	Male	Left temporal	114.7	0	Metastasis	8 | 4	0.3
24	Female	Left frontal	0.2	0	Hemangioblastoma	5 | 5	0.5
74	Male	Right frontal	29.3	0	Meningioma	6 | 5	0
54	Female	Left occipital	3.6	0	Glioblastoma	5 | 6	0.5
37	Female	Right frontal	161.0	0	Meningioma	7 | 4	0

* Zero indicates that tumor reached brain surface.

Discussion

This prospective clinical study offers a proof of concept of the clinical feasibility of a wearable mixed reality device for pre-operative neurosurgical planning using 3D hologram reconstructions of brain tumors with quantitative outcome measures. This study shows that there was a median deviation of 0.4 cm in tumor localization, when the Hololens was compared with the standard neuronavigation system. In nine out of 25 patients the localization did not differ between both systems. Additionally, the measurement strategy used in this study, focusing on the tumor center, was qualitatively assessed for full tumor alignment and showed that the full tumor alignment on the patient's head using the Hololens corresponds with the true tumor borders as seen on the brain surface.

Surgeons experienced benefits in terms of preservation of attention and focus on the patient, improved ergonomics, and an improved understanding of tumor-brain/skull relationship due to a direct 3D holographic representation. To the best of our knowledge, there have been no studies on wearable mixed reality devices with 3D virtual projections for brain tumor surgery to this date. We emphasize that spatial mapping and recognition of the physical environment is a crucial difference between augmented and mixed reality. Mixed reality with the Hololens provided an instant visualization of tumors by using 3D holograms. Due to a built-in computer in the Hololens, other devices such as camera, monitor or excessive technical preparation at the operating room were not necessary. In contrast to current neuronavigation systems, the Hololens subjectively improved ergonomics during surgical planning, because neurosurgeons could keep their focus on patient's head, instead of having to look back and forth between patient and navigation screen. Importantly, during this study, there was a learning curve for neurosurgeons in using this wearable mixed reality device in localizing and marking tumor border on the patient's head as evidenced by a trend towards more accurately localizing tumors towards the end of the study.

Within the field of neurosurgery, VR has played a limited role in pre- and intra-operative application due to the fact that the neurosurgeon is isolated in a virtual world, while being in the operating room were situational awareness is crucial. However, for extra-operative pre-surgical planning and education, VR devices are useful to improve the understanding of both individual anatomy of patients [7], anatomical education [8] and neurosurgical traning[5]. Meola et al. reviewed 18 articles on AR in neurosurgery [6] published between 1996 and 2015. Microscopes [9-16] and monitors or tablets [17-24] were the most commonly used AR devices, which were all 2D visualizations of virtual objects such as aneurysms and brain tumors. Monitors or tablets could be practical AR devices to use for pre-operative surgical planning. However, with these devices, an in-physical-space 3D visualization of virtual objects of complex anatomical structures is limited. Additionally, microscopes offer some 3D experience of virtual objects such as tumor, fiber tracts and

vessels trough the oculars. However, the microscope is commonly used after trepanation in the intra-operative phase, instead of a pre-operative planning device.

Recently, Yoon et al. reviewed the literature between 1995 and 2017 extensively on several augmented reality devices within all surgical fields.[25] Within the field of neurosurgery, they identified 10 studies, of which 5 performed surgery in a live setting and 5 performed simulated surgeries. There were no studies that had used the Hololens. The authors found that wearable augmented reality devices improved attention, focus and ergonomics compared to standard navigation systems in which the surgeon have to shift repeatedly from 2D images on a screen to the patient. The authors also state that the majority of the studies had not assessed any quantitative outcome measure. The experience and results in our study concerning the preservation of attention and focus by the surgeon on the patient, the improved 3D anatomical evaluation, and improved ergonomics are in line with the conclusions of this systematic review.

This study provides a proof of concept that the Hololens has potential for operative planning of brain tumor surgery with quantitative outcome measures. However, we think that there are two important issues that need further development to improve the accuracy of the Hololens. Due to the manual skin surface registration, the accuracy was probably decreased and registration time longer as compared to the standard neuronavigation. The Hololens deviated with a median of 0.4 cm in localizing the tumor and prolonged median registration time with 45 seconds. In the setting of operative planning, this deviation is probably acceptable, it is however too large for intra-operative usage. Currently, we are working on a coordination-based skin surface registration system which is expected to improve spatial accuracy and decrease registration time. Additionally, there was an issue with respect to the point of view when marking tumor border on the skin of the patient's head with the Hololens, when compared with standard neuronavigation. Since the holographic head is 3D and transparent, the tumor can be seen and marked from different angles. Any angle that is not exactly perpendicular to the head and tumor, will result in a deviation of localizing the lesion when compared with standard neuronavigation.

A limitation of this study is that a non-validated method was used in this study to outline the tumor. This method is standard practice with the use of the neuronavigation system in our neurosurgical center, and we therefore chose to use this method also in context of this study for the Hololens. Another limitation of this study is that patients were not consecutively included in this study, which could have introduced selection bias based on case complexity. Most of the tumors were superficially located. However, we did have some complex cases with tumors in the basal ganglia or skull base. We did not find any significant difference in accuracy between superficial and deep tumors. In cases with skull base tumors, surgeons reported a better understanding of tumor-brain/skull base relationship.

Further research is needed to test the applicability and accuracy of mixed reality devices in several other neurosurgical fields such as pre and per-operative planning of

neurosurgery by visualizing neuroanatomical structures (such as more complex tumors, aneurysms, vascular malformations, skull base tumors, and disc herniations). Additional options of this technology should be explored, which includes overlaying patient (history, vital statistics) and imaging data (MRI and/or CT scan) directly over the field of view of the surgeon, which could further improve interpretation, attention and focus of the surgeon during surgery.

Conclusion

This prospective clinical study offers a proof of concept of the clinical feasibility of the Hololens for brain tumor surgery planning in the operating room with quantitative outcome measures. Further development is needed to improve the accuracy and clinical applicability of this wearable mixed reality device.

References

1. Wadley J, Dorward N, Kitchen N, Thomas D. Pre-operative planning and intra-operative guidance in modern neurosurgery: a review of 300 cases. ***Ann R Coll Surg Engl.*** Jul 1999;81(4):217-225.
2. Barone DG, Lawrie TA, Hart MG. Image guided surgery for the resection of brain tumours. ***Cochrane Database Syst Rev.*** Jan 28 2014(1):CD009685.
3. Willems PW, Taphoorn MJ, Burger H, Berkelbach van der Sprenkel JW, Tulleken CA. Effectiveness of neuronavigation in resecting solitary intracerebral contrast-enhancing tumors: a randomized controlled trial. ***J Neurosurg.*** Mar 2006;104(3):360-368.
4. Kishino Ma. A Taxonomy of Mixed Reality Visual Displays. ***IEICE Transactions on Information and Systems*** 1994;vol. E77-D, no. 12(12):1321-1329
5. Pelargos PE, Nagasawa DT, Lagman C, et al. Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery. ***J Clin Neurosci.*** Jan 2017;35:1-4.
6. Meola A, Cutolo F, Carbone M, Cagnazzo F, Ferrari M, Ferrari V. Augmented reality in neurosurgery: a systematic review. ***Neurosurg Rev.*** May 07 2016.
7. Kockro RA, Killeen T, Ayyad A, et al. Aneurysm Surgery with Preoperative Three-Dimensional Planning in a Virtual Reality Environment: Technique and Outcome Analysis. ***World Neurosurg.*** Dec 2016;96:489-499.
8. de Faria JW, Teixeira MJ, de Moura Sousa Junior L, Otoch JP, Figueiredo EG. Virtual and stereoscopic anatomy: when virtual reality meets medical education. ***J Neurosurg.*** Nov 2016;125(5):1105-1111.
9. Sun GC, Wang F, Chen XL, et al. Impact of Virtual and Augmented Reality Based on Intraoperative Magnetic Resonance Imaging and Functional Neuronavigation in Glioma Surgery Involving Eloquent Areas. ***World Neurosurg.*** Dec 2016;96:375-382.
10. Cabrilo I, Bijlenga P, Schaller K. Augmented reality in the surgery of cerebral aneurysms: a technical report. ***Neurosurgery.*** Jun 2014;10 Suppl 2:252-260; discussion 260-251.
11. Cabrilo I, Bijlenga P, Schaller K. Augmented reality in the surgery of cerebral arteriovenous malformations: technique assessment and considerations. ***Acta Neurochir (Wien).*** Sep 2014;156(9):1769-1774.
12. Cabrilo I, Schaller K, Bijlenga P. Augmented reality-assisted bypass surgery: embracing minimal invasiveness. ***World Neurosurg.*** Apr 2015;83(4):596-602.
13. Edwards PJ JL, Hawkes DJ, Fenlon MR, Strong A,, M G. Clinical experience and perception in stereo augmented reality surgical navigation. ***MIAR. Springer-Verlag, Berlin,*** *pp 369-376*. 2004.
14. King AP, Edwards PJ, Maurer CR, Jr., et al. A system for microscope-assisted guided interventions. ***Stereotact Funct Neurosurg.*** 1999;72(2-4):107-111.
15. Stadie AT, Reisch R, Kockro RA, et al. Minimally invasive cerebral cavernoma surgery using keyhole approaches - solutions for technique-related limitations. ***Minim Invasive Neurosurg.*** Feb 2009;52(1):9-16.

16. Sven R. Kantelhardt AG, Axel Neulen, Naureen Keric, Mirjam Renovanz, Alf Giese. Video-Assisted Navigation for Adjustment of Image-Guidance Accuracy to Slight Brain Shift. 2015. ***Oper Neurosurg (Hagerstown).*** 2015 Dec 1;11(4):504-511.
17. Deng W, Li F, Wang M, Song Z. Easy-to-use augmented reality neuronavigation using a wireless tablet PC. ***Stereotact Funct Neurosurg.*** 2014;92(1):17-24.
18. Inoue D, Cho B, Mori M, et al. Preliminary study on the clinical application of augmented reality neuronavigation. ***J Neurol Surg A Cent Eur Neurosurg.*** Mar 2013;74(2):71-76.
19. Kawamata T, Iseki H, Shibasaki T, Hori T. Endoscopic augmented reality navigation system for endonasal transsphenoidal surgery to treat pituitary tumors: technical note. ***Neurosurgery.*** Jun 2002;50(6):1393-1397.
20. Kockro RA, Tsai YT, Ng I, et al. Dex-ray: augmented reality neurosurgical navigation with a handheld video probe. ***Neurosurgery.*** Oct 2009;65(4):795-807; discussion 807-798.
21. Lovo EE, Quintana JC, Puebla MC, et al. A novel, inexpensive method of image coregistration for applications in image-guided surgery using augmented reality. ***Neurosurgery.*** Apr 2007;60(4 Suppl 2):366-371; discussion 371-362.
22. Low D, Lee CK, Dip LL, Ng WH, Ang BT, Ng I. Augmented reality neurosurgical planning and navigation for surgical excision of parasagittal, falcine and convexity meningiomas. ***Br J Neurosurg.*** Feb 2010;24(1):69-74.
23. Masutani Y, Dohi T, Yamane F, Iseki H, Takakura K. Augmented reality visualization system for intravascular neurosurgery. ***Comput Aided Surg.*** 1998;3(5):239-247.
24. Paul P, Fleig O, Jannin P. Augmented virtuality based on stereoscopic reconstruction in multimodal image-guided neurosurgery: methods and performance evaluation. ***IEEE Trans Med Imaging.*** Nov 2005;24(11):1500-1511.
25. Yoon W, Chen R, Kim E, Akinduro O, Kerezoudis P, Han P et al. Augmented reality for the surgeon: Systematic review. Int J Med Robotics Comput Assist Surg. 2018;1–13.

PART III

Extent of resection and survival

Chapter 8

The association between the extent of glioblastoma resection and survival in light of MGMT promoter methylation in 326 patients with newly diagnosed IDH wildtype glioblastoma

Fatih Incekara, Marion Smits, Sebastian R. van der Voort, Hendrik Jan Dubbink, Peggy N. Atmodimedjo, Johan M. Kros, Arnaud J.P.E. Vincent, Martin van den Bent

Frontiers in Oncology. 2020 Jul 10;10:1087

Abstract

Background

Studies on the impact of contrast enhanced (CE) and non-contrast enhanced (NCE) tumor resection in patients with glioblastoma in light of molecular subtypes are limited. The aim of this study was to assess the impact of CE and NCE tumor resection in light of MGMT promoter methylation on survival in newly diagnosed IDH wildtype glioblastoma.

Methods

Patients with newly diagnosed IDH wildtype glioblastoma who underwent surgery were eligible. CE and NCE tumor volumes were assessed on pre- and post-operative MRI scans and extent of resection was calculated. The impact of CE and NCE resection was evaluated using multivariable Cox proportional hazards models and Kaplan Meier analyses.

Results

326 patients were included: 177 (54.3%) with and 149 (45.7%) without MGMT methylation. Multivariable Cox proportional hazards models stratified for MGMT methylation identified age ≤ 65y (HR 0.63; 95% CI, 0.49-0.81; $p < 0.0001$), chemoradiation (HR 0.13; 95% CI, 0.09-0.19; $p < 0.0001$), maximal CE resection (HR 0.58; 95% CI, 0.39-0.87; $p = 0.009$), extended NCE resection (HR 0.71; 95% CI, 0.53-0.93; $p = 0.014$) and minimal residual CE tumor volume (HR 0.64; 95% CI, 0.46-0.88 $p = 0.007$) as being associated with longer OS. Kaplan Meier analyses showed that extensive surgery was more beneficial for patients with MGMT methylated glioblastoma.

Conclusion

This study shows an association between maximal CE resection, extended NCE resection, minimal residual CE tumor volume and longer OS in patients with newly diagnosed IDH wildtype glioblastoma. Intra-operative imaging and stimulation mapping may be used to pursue safe and maximal resection.

Introduction

Patients with glioblastoma have a poor prognosis with a median overall survival of 10 - 15 months, despite safe and maximal surgical resection followed by chemo- and radiotherapy.[1] This prognosis varies based on known factors such as age, KPS, extent of resection, isocitrate dehydrogenase (IDH) mutation status, and methylguanine methyltransferase (MGMT) promoter methylation status.[2,3]

Maximal resection of the contrast enhanced (CE) portion of glioblastoma has been associated with better overall survival and is currently part of standard surgical glioblastoma treatment.[3] However, glioblastoma is known to infiltrate far beyond the margins of CE as seen on MRI, into the surrounding non-contrast enhanced (NCE) edematous T2-weighted or FLAIR abnormality area.[4] This raises the question whether maximal CE resection should be extended beyond CE, into NCE area, to improve survival.[5] A recent meta-analysis and a systematic review suggested that there is an association between maximal CE resection with resection of NCE and overall survival.[6,7] However, the quality of evidence of the available studies was low due to confounding and selection biases. On top, studies investigating the impact of CE and NCE resection have reported limited molecular data on IDH mutation and MGMT promoter methylation of their studied glioblastoma population.[6,8]

Thus, in light of the WHO 2016 reclassification, which now includes such molecular data, the impact of CE and NCE glioblastoma resection needs to be re-evaluated in a molecularly homogenous glioblastoma IDH wildtype population, while considering the impact of MGMT promoter methylation.[9] The aim of this study therefore, was to assess the impact of CE and NCE tumor resection in light of MGMT promoter methylation on survival in a cohort of patients with newly diagnosed IDH wildtype glioblastoma.

Methods

Patients

All consecutive patients aged 18 years or older, newly diagnosed with a CE mass lesion as seen on post-contrast T1-weighted MRI scans, histopathological confirmed as IDH wildtype glioblastoma, who underwent tumor resection or biopsy between January 2012 and May 2018 at […] were considered for this retrospective study. Patients were eligible if pre- and immediate post-operative (<48 hours) T2-weighted or FLAIR and post-contrast T1-weighted MRI scans were available together with complete molecular data on IDH mutation and MGMT methylation. Molecular analysis was post-hoc performed in patients with unknown IDH mutation or MGMT methylation status; patients without enough tumor material for molecular analysis or in whom assays failed to produce a test result were excluded. The study was approved by the Medical Ethical Committee of Erasmus MC, who waived the need for written informed consent from the patients due to the retrospective

nature of this study and the (emotional) burden that would result from contacting the patients or their relatives to obtain consent. The study was performed in accordance with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards and reported following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.

Image acquisition, tumor segmentation and extent of resection

From pre- and post-operative MRI scans, which were obtained in the clinical routine either on a 1.5T or 3.0T scanner, post-contrast T1-weighted and T2-weighted or FLAIR images were collected.

For glioblastoma segmentation, we imported both pre- and post-operative post-contrast T1-weighted and T2-weighted or FLAIR scans into Brainlab (BrainLab, Feldkirchen, Germany, version 2.1.0.15). Using the SmartBrush tool in Brainlab Elements, we semi-automatically segmented all tumor involved CE on pre-operative post-contrast T1-weighted scans (including the necrotic part, if present) and all tumor involved CE on post-operative post-contrast T1-weighted scans (excluding small vessels in the surgical cavity or hemorrhage). We then semi-automatically segmented all tumor-related NCE on both pre- and post-operative T2-weighted or FLAIR scans (excluding extra-lesional hemorrhage). We attempted to minimize the inclusion of surgery induced new T2-weighted or FLAIR abnormality by overlaying and carefully comparing pre- and post-operative MRIs. We manually corrected all segmentations when needed using the manual Brush tool. All tumor volumes were assessed while being blinded for patients' clinical outcome.

We finally obtained four tumor volumes (cm^3): pre-operative and residual CE volumes and pre-operative and residual NCE volumes. We calculated the CE surrounding NCE volumes by subtracting CE volumes from the total NCE volumes. We calculated the extent of resection (EOR, %) separately for both the CE and NCE portion with the formula: [(pre-operative volume – residual volume)/pre-operative volume] * 100.[10] Maximal CE resection was categorized in our dataset as CE EOR >97% and extended NCE was categorized as NCE resection of ≥ 30% based on threshold analysis (Figure 1 and 2 of Supplementary Material). Tumors that were biopsied were segmented only on pre-operative MRI scans and their EOR was imputed as being 0%.

Molecular analysis

Tumor tissue samples were obtained from patients through surgical resection or biopsy. Histopathological examination was performed by neuropathologists; IDH mutational analysis was assessed with sequencing and MGMT methylation status with a methylation specific PCR, as described elsewhere.[11-13]

Statistical analysis

Overall survival (OS) was defined as time from surgery to death (primary outcome) and progression-free survival (PFS) was defined as time from surgery till clinical or radiological progression (secondary outcome). Patients were censored at time of last clinical follow up date.

Statistical analyses were performed with SPSS 25.0 statistical software (IBM Corp.). Pre-operative and post-operative residual tumor volume distributions were skewed and therefore log transformed prior to statistical analysis. Descriptive statistics were tested between MGMT methylation status groups with the Chi Squared Test or Fisher Exact test in case of categorical variables, with the Kruskal Wallis test in case of continuous non-normal distributed data and with log rank tests to compare median OS and PFS when using Kaplan Meier analysis.

The associations between each variable and outcome were first tested with univariable Cox proportional hazards models and all variables with $p < 0.10$ (entry significance threshold) were selected for multivariable Cox proportional hazards models. These models were stratified for MGMT, because this variable violated the proportional hazards assumption. Hazard Ratio's (HR) with 95% Confidence Intervals (CI) were estimated for each variable within the model. A p-value < 0.05 was considered statistically significant.

Results

There were 375 glioblastoma patients considered for this study. We excluded 36 (9.6%) due to insufficient tissue material for molecular analysis, and 13 (3.5%) because of the presence of IDH mutation. In total, 326 IDH wildtype glioblastoma were included in our analysis: 177 (54.3%) with and 149 (45.7%) without MGMT promoter methylation. Maximal CE resection was achieved in 61 patients (18.7%), while in 187 (57.4%) patients maximal resection of CE could not be achieved. Seventy-eight patients (23.9%) underwent biopsy. Extended NCE was achieved in 156 patients (47.9%) and no or limited NCE resection was performed in 170 patients (52.1%). Median OS and PFS was 309 days (95% CI, 278-340) and 174 days (95% CI, 159-209) respectively. Further patient and tumor characteristics are presented in Table 1.

Table 1. Patient and tumor characteristics.

	MGMT promoter						
	All		**Methylated**		**Unmethylated**		**p value**
	n	**%**	**n**	**%**	**n**	**%**	
Characteristics	326	100	177	54.3	149	45.7	
Sex							**0.006**
Male	206	63.2	100	48.5	106	51.5	
Female	120	36.8	77	64.2	43	35.8	
Age, years							0.122
≤ 65	162	49.7	81	50.0	81	50.0	
> 65	164	50.3	96	58.5	68	41.5	
Mean, years (SD)	63.8 (10.5)		64.2 (10.9)		63.3 (10.1)		0.453
KPS							0.748
≤ 70	119	36.5	66	55.5	53	45.5	
> 70	207	73.5	111	53.6	96	46.4	
Mean (SD)	79.2 (12.4)		78.8 (12.2)		79.7 (12.6)		0.480
Pre-operative tumor volume, median cm^3 (IQR)							
CE	34.9 (15.5-55.8)		30.7 (13.8-52.5)		39.9 (17.1-26.2)		0.050
NCE	72.1 (29.4-127.5)		78.1 (27.8-133.8)		64.9 (32.3-113.4)		0.249
Residual tumor volume, median cm^3 (IQR)							
CE	5.0 (1.41-12.3)		6.1 (1.8-12.5)		3.8 (1.1-12.0)		0.097
NCE	39.7 (18.3-73.7)		42.2 (17.1-79.6)		36.5 (20.4-64.2)		0.699
Maximal CE resection							0.213
Yes	61	18.7	27	44.3	34	55.7	
No	187	57.4	105	56.2	82	43.8	
Biopsy	78	23.9	45	57.7	33	42.3	
Median EOR (IQR)	83.4 (14.2-94.8)		79.6 (0-91.9)		88.0 (25-7-96.0)		0.075
NCE resection							0.609
≥ 30%	156	47.9	87	55.8	69	44.2	
< 30%	170	52.1	90	52.9	80	47.1	
Median EOR (IQR)	27.1 (0-57.1)		28.7 (0-58.4)		25.8 (0-55.0)		0.948
Adjuvant therapy							0.294
No therapy	61	18.7	38	62.3	23	37.7	
Radio- or chemotherapy alone	56	17.2	27	48.2	29	51.8	
Chemoradiation	209	64.1	112	58.4	97	41.6	
Overall patient outcome	x	x	x	x	x	x	x
Median overall survival (95% CI)	309 (278.0-340.0)		334 (266.8-401.2)		305 (275.2-334.8)		**0.003**
Median progression-free survival (95% CI)	184 (159.3-208.7)		174 (131-2-216.8)		190 (162.0-218)		0.053

MGMT methylguanine methyltransferase, SD standard deviation, CE contrast enhancement, NCE non contrast enhancement, EOR extent of resection, IQR Inter Quartile Range, CI Confidence interval.

Univariable Cox proportional hazards regression analysis identified age ≤ 65y (HR 0.64; 95% CI, 0.51-0.80; p < 0.0001), KPS >70 (HR 0.59; 95% CI, 0.47-0.74; p < 0.0001), MGMT promoter methylation (HR 0.72; 95% CI, 0.57-0.90; p < 0.004), adjuvant chemoradiation (HR 0.14; 95% CI, 0.11-0.19; p < 0.0001), smaller pre-operative CE tumor volumes (per cm^3 HR 0.92; 95% CI, 0.83-1.01; p < .094), maximal CE resection (HR 0.51; 95% CI, 0.36-0.72; p < 0.0001), and extended NCE resection (HR 0.72; 95% CI, 0.58-0.91; p = 0.005) as being associated with longer OS. As further presented in Table 2, the variables age ≤ 65y, KPS >70, adjuvant chemoradiation, maximal CE resection and extended NCE resection were also significantly associated with a longer PFS in univariable Cox regression analysis (p < 0.05). Kaplan Meier curves for OS and PFS for each variable are presented in Figure 3 of Supplementary Material.

Multivariable Cox proportional hazards regression analysis stratified for MGMT methylation status and risk adjusted for age ≤ 65y (HR 0.63; 95% CI, 0.49-0.81; p < 0.0001), KPS >70 (HR 0.92; 95% CI, 0.71-1.19; p = 0.545), adjuvant chemoradiation (HR 0.13; 95% CI, 0.09-0.19; p < 0.0001), and smaller pre-operative CE tumor volumes per cm^3 (HR 0.85; 95% CI, 0.75-0.95; p = 0.007) identified maximal CE resection (HR 0.58; 95% CI, 0.39-0.87; p = 0.009) and extended NCE resection (HR 0.71; 95% CI, 0.53-0.93; p = 0.014) as being associated with longer OS (Table 2). Variables that remained significantly associated with a longer PFS were KPS >70 (HR 0.59; 95% CI, 0.45-0.77; p < 0.0001) and adjuvant chemoradiation (HR 0.09; 95% CI, 0.06-0.14; p < 0.0001). Explorative multivariable Cox proportional hazards regression analyses showed that higher NCE resection thresholds (e.g. ≥50%) were not associated with a favorable OS (p > 0.05) (for threshold analysis, see Figure 1 of Supplementary Material).

The impact of maximal CE resection on survival was more beneficial for patients with MGMT methylated glioblastoma and significantly improved median OS (572 days; 95% CI, 424-720), when compared to STR (342 days; 95% CI, 282-402; p = 0.014) or biopsy (112 days; 95% CI, 36-42; p = 0.001) (Figure 1A). Patients with MGMT methylated glioblastoma also had a longer OS with extended NCE resection (425 days; 95% CI, 286-564) than when extended NCE was not achieved (190 days; 95% CI, 107-273; p = 0.001) (Figure 1B). In patients with MGMT unmethylated glioblastoma, no survival benefit was observed with extended NCE resection (p = 0.884).

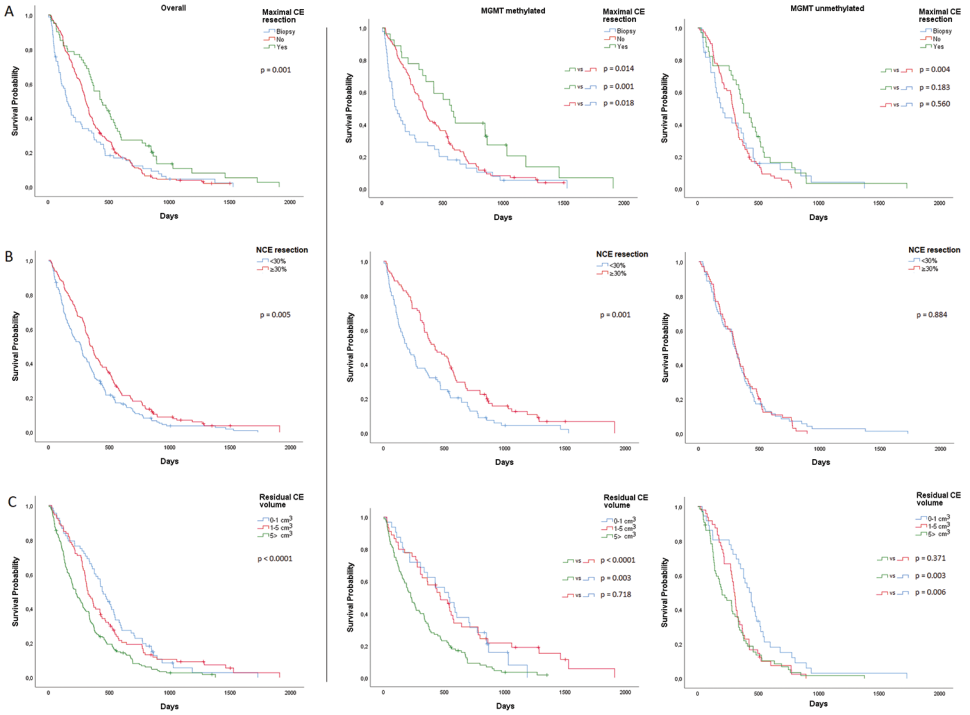


Figure 1. Overall and MGMT-stratified Kaplan Meier curves for overall survival A) Maximal CE resection B) Extended NCE resection and C) Minimal residual volume.

We further assessed minimal post-operative CE residual volumes with a potential positive impact on OS. In MGMT methylated glioblastoma, when compared to >5 cm^3 residual CE tumor (224 days; 95% CI 164-284), we observed a significantly longer median OS for a residual CE tumor volume of 1-5 cm^3 (470 days; 95% CI 330-610; p <0.0001) and 0-1 cm^3 (536 days; 95% CI 319-752; p = 0.003) (Figure 1C). In MGMT unmethylated glioblastoma, we only observed a longer OS in residual CE volumes of 0-1 cm^3 (427 days; 95% CI, 350-503), when compared to 1-5cm^3 (299 days; 95% CI, 270-327; p = 0.006) or to >5 cm^3 (200 days; 95% CI, 102-298; p = 0.003) (Figure 1C). In these tumors, no difference in median OS was observed between 1-5 cm^3 and >5 cm^3 (p = 0.371) residual CE tumor volume. Multivariable Cox proportional hazards regression analysis stratified for MGMT and adjusted for age, KPS, adjuvant chemoradiation and pre-operative CE tumor volume, identified a residual CE tumor volume of 0-1 cm^3 (HR 0.64; 95% CI, 0.46-0.88 p = 0.007) and 1-5 cm^3 (HR 0.71; 95% CI, 0.54-0.94; p = 0.016) as being associated with favorable OS. These analyses were also performed for post-operative residual NCE volumes, but here no minimal volume threshold with a positive impact on OS was identified.

Table 2. Cox proportional hazards models for overall survival and progression-free survival.

	Overall survival					Progression-free survival				
		Univariable		Multivariable			Univariable		Multivariable	
Variable	**Median days, 95%CI**	**HR, 95% CI**	**P Value #**	**Adjusted HR, 95% CI**	**P Value §**	**Median days, 95%CI**	**HR, 95% CI**	**P Value #**	**Adjusted HR, 95% CI**	**P Value §**
Age, years										
> 65	259, 191-327	a				144, 114-175	a			
≤ 65	358, 311-405	0.64, 0.51-0.80	< **0.0001**	0.63, 0.49-0.81	< **0.0001**	216, 181-251	0.70, 0.56-0.89	**0.003**	0.82, 0.64-1.06	0.123
KPS										
≤ 70	199, 142-256	a				120, 91-149	a			
> 70	369, 319-419	0.59, 0.47-0.74	< **0.0001**	0.92, 0.71-1.19	0.545	223, 181-265	0.51, 0.40-0.6	< **0.0001**	0.59, 0.45-0.77	< **0.0001**
MGMT promoter										
Unmethylated	305, 275-335	a		b		190, 162-218	a		b	
Methylated	334, 267-401	0.72, 0.57-0.90	**0.004**			174, 131-217	0.79, 0.62-1.00	**0.054**		
Adjuvant therapy										
None	65, 46-84	a				47, 38-56	a			
Radiotherapy or chemotherapy	217, 146-288	0.38, 0.26-0.55	< **0.0001**	0.29, 0.19-0.43	< **0.0001**	129, 107-151	0.23, 0.15-0.36	< **0.0001**	0.22, 0.14-0.35	< **0.0001**
Chemoradiation	423, 367-479	0.14, 0.11-0.19	< **0.0001**	0.13, 0.09-0.19	< **0.0001**	254, 222-287	0.09, 0.06-0.13	< **0.0001**	0.09, 0.06-0.14	< **0.0001**
Pre-operative CE tumor volume										
smaller per cm^3		0.92, 0.83-1.01	**0.094**	0.85, 0.75-0.95	**0.007**		0.94, 0.85-1.04	0.239	0.99, 0.88-1.11	0.860
Maximal CE resection	*					*				
Biopsy	153, 106-200	a				100, 87-113	a			
No	309, 278-340	0.79, 0.60-1.04	**0.090**	0.87, 0.61-1.23	0.415	190, 170-210	0.80, 0.61-1.05	0.109	1.25, 0.89-1.81	0.228
Yes	441, 321-560	0.51, 0.36-0.72	< **0.0001**	0.58, 0.39-0.87	**0.009**	272, 210-334	0.59, 0.42-0.85	**0.004**	0.97, 0.63-1.50	0.898
NCE resection	*					*				
< 30%	266, 213-319	a				155, 104-206	a			
≥ 30%	344, 304-385	0.72, 0.58-0.91	**0.005**	0.71, 0.53-0.93	**0.014**	203, 166-240	0.82, 0.65-1.03	**0.087**	0.87, 0.65-1.15	0.325

[a] Reference group [b] Proportional hazard assumption for MGMT was violated, therefore multivariable Cox regression model was stratified for MGMT. HR Hazard ratio, MGMT methylguanine methyltransferase, CI Confidence Interval *for MGMT-stratified median overall survival, see Figure 1 # entry significance threshold of $p < 0.1$ § significance threshold of $p < 0.05$

Discussion

This study shows an association between maximal CE resection, extended NCE resection, minimal residual CE tumor volume and longer OS in 326 patients with newly diagnosed IDH wildtype glioblastoma. We observed that extensive resection was more beneficial for patients with MGMT methylated IDH wildtype glioblastoma.

Maximal resection of CE has earlier been associated with longer OS in a large meta-analysis of neurosurgical literature based on 37 studies and 41,117 unique glioblastoma patients.[3] This association has recently been re-evaluated based on two new insights. First, studies performed before the WHO 2016 reclassification have included limited molecular data, because the impact of molecular subtyping of glioblastoma according to IDH mutation status was less of a consideration.[3,8,14-16] Although IDH mutation within newly diagnosed primary glioblastoma is rare (<5%, and 3.5% in our cohort) these tumors represent a distinct molecular type of glioma arising from a distinct precursor lesion.[17,18] Therefore, incomplete or absence of molecular data on IDH mutation and MGMT methylation or mixing molecular subtypes when evaluating the impact of glioblastoma resection on survival is undesirable. More recent studies did investigate the impact of glioblastoma surgery on survival in light of molecular markers. Senft et al. studied the impact of maximal CE resection and MGMT promoter methylation status in a homogenous IDH wildtype glioblastoma population (n = 175) and showed that both were significantly associated with longer OS.[2] Ellingson et al. showed in 1,054 glioblastoma patients (with partially available data on IDH mutation and MGMT methylation) that smaller residual CE tumor volumes (<12 cm^3) and MGMT methylation were significantly associated with longer OS in patients receiving chemoradiation.[19] A recent study published by Molinaro et al confirmed the association between maximal CE resection and OS across all molecular subgroups of glioblastoma.[20]

Secondly, the association between glioblastoma resection and OS is also being reassessed by evaluating the value of NCE resection, because it is known that glioblastoma infiltrates beyond the margins of CE into the NCE area.[4] This aspect of glioblastoma surgery is also recently investigated by Molinaro et al.[20] The authors found that maximal resection of CE and NCE tumor was associated with longer OS in younger patients with IDH wildtype glioblastoma regardless of MGMT methylation status (subset of 190 patients with known IDH mutation and MGMT methylation status). In this study, maximal NCE resection that was associated with OS was defined as 92% NCE resection after maximal CE resection. Such an extensive surgical approach may only safely be achievable with use of intra-operative imaging, fluorescent guidance or stimulation mapping.[21,22] Other studies have associated lower NCE resection thresholds of 53% and 45% with OS.[8,23] We observed that a NCE resection threshold of 30% was associated with OS. In exploratory threshold analysis, higher thresholds (e.g. minimal ≥60% NCE tumor resection) seemed not to be associated with OS anymore. This may suggest that resection of NCE tumor immediately

surrounding CE improves survival, but extending the resection on further distance from CE into NCE tumor does not provide survival benefit. It can be hypothesized that the direct peritumoral NCE area reflects a higher degree of tumor infiltration than the NCE area further away from the CE tumor, which is presumably more dominated by edema than tumor infiltration.[4] In future research, a combination of physiological imaging modalities - such as MR spectroscopy, diffusion and perfusion imaging or positron-emission tomography - may be used to more accurately detect tumor infiltrated portions in NCE and to tailor surgical planning.[24]

Limitations

The main limitation of this study is its retrospective nature. This may have introduced some degree of selection bias. We attempted to limit selection bias by consecutive inclusion of all glioblastoma patients operated upon between 2012 and 2018 in our cohort, including patients with complex glioblastoma localization (crossing midline or deep within the basal ganglia) who underwent diagnostic biopsies. We also performed IDH mutational and MGMT promoter methylation analyses on all glioblastoma included in our cohort. A second limitation is that only one observer assessed both pre- and post-operative tumor volumes. In this context, a stringent assessment of residual volumes in the resection cavity by one observer may explain the relatively low maximal CE resection percentage of 17.8%. Although the interobserver agreement is high for pre-operative volumes, it is known to be relatively low for residual tumor volumes. The intra-observer agreement nevertheless, is known to be high for both pre-operative as residual tumor volumes.[10,25] We have also attempted to limit bias during volumetric assessment by blinding the assessor for patients' clinical outcome. In future research, our findings need to be validated in an external validation set.

Conclusion

This study shows an association between maximal CE resection, extended NCE resection, minimal residual CE tumor volume and longer OS in patients with newly diagnosed IDH wildtype glioblastoma. Extensive resection was more beneficial for patients with MGMT methylated glioblastoma. Intra operative imaging and stimulation mapping may be used to pursue maximal CE resection and extended NCE resection.

References

1. Stupp R, Mason WP, Van Den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New Engl J Med.* 2005;352(10):987-996.
2. Gessler F, Bernstock JD, Braczynski A, et al. Surgery for Glioblastoma in Light of Molecular Markers: Impact of Resection and MGMT Promoter Methylation in Newly Diagnosed IDH-1 Wild-Type Glioblastomas. *Neurosurgery.* Jan 1 2019;84(1):190-197.
3. Brown TJ, Brennan MC, Li M, et al. Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis. *JAMA Oncol.* 2016;2(11):1460-1469.
4. Eidel O, Burth S, Neumann JO, et al. Tumor Infiltration in Enhancing and Non-Enhancing Parts of Glioblastoma: A Correlation with Histopathology. *PLoS One.* 2017;12(1):e0169292.
5. de Leeuw CN, Vogelbaum MA. Supratotal resection in glioma: a systematic review. *Neuro Oncol.* Feb 14 2019;21(2):179-188.
6. Incekara F, Koene S, Vincent A, van den Bent MJ, Smits M. Association Between Supratotal Glioblastoma Resection and Patient Survival: A Systematic Review and Meta-Analysis. *World Neurosurg.* Jul 2019;127:617-624 e612.
7. de Leeuw CN, Vogelbaum MA. Supratotal resection in glioma: a systematic review. *Neuro Oncol.* 2019;21(2):179-188.
8. Li YM, Suki D, Hess K, Sawaya R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? *J Neurosurg.* 2016;124(4):977-988.
9. 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.* Jun 2016;131(6):803-820.
10. Grabowski MM, Recinos PF, Nowacki AS, et al. Residual tumor volume versus extent of resection: Predictors of survival after surgery for glioblastoma. *J Neurosurg.* 2014;121(5):1115-1123.
11. Taal W, Oosterkamp HM, Walenkamp AM, et al. Single-agent bevacizumab or lomustine versus a combination of bevacizumab plus lomustine in patients with recurrent glioblastoma (BELOB trial): a randomised controlled phase 2 trial. *Lancet Oncol.* Aug 2014;15(9):943-953.
12. van der Voort SR, Incekara F, Wijnenga MM, et al. Predicting the 1p/19q co-deletion status of presumed low grade glioma with an externally validated machine learning algorithm. *Clin Cancer Res.* Sep 23 2019.
13. Dubbink HJ, Atmodimedjo PN, Kros JM, et al. Molecular classification of anaplastic oligodendroglioma using next-generation sequencing: a report of the prospective randomized EORTC Brain Tumor Group 26951 phase III trial. *Neuro Oncol.* Mar 2016;18(3):388-400.
14. Stummer W, Pichlmeier U, Meinel T, Wiestler OD. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. *The lancet oncology.* 2006.
15. Senft C, Bink A, Franz K, Vatter H, Gasser T, Seifert V. Intraoperative MRI guidance and extent of resection in glioma surgery: a randomised, controlled trial. *Lancet Oncol.* Oct 2011;12(11):997-1003.

16. Lacroix M, Abi-Said D, Fourney DR, et al. A multivariate analysis of 416 patients with glioblastoma multiforme: Prognosis, extent of resection, and survival. *J Neurosurg.* 2001;95(2):190-198.
17. Lai A, Kharbanda S, Pope WB, et al. Evidence for sequenced molecular evolution of IDH1 mutant glioblastoma from a distinct cell of origin. *J Clin Oncol.* Dec 1 2011;29(34):4482-4490.
18. Yan H, Parsons DW, Jin G, et al. IDH1 and IDH2 mutations in gliomas. *N Engl J Med.* Feb 19 2009;360(8):765-773.
19. Ellingson BM, Abrey LE, Nelson SJ, et al. Validation of postoperative residual contrast-enhancing tumor volume as an independent prognostic factor for overall survival in newly diagnosed glioblastoma. *Neuro Oncol.* Aug 2 2018;20(9):1240-1250.
20. Molinaro AM, Hervey-Jumper S, Morshed RA, et al. Association of Maximal Extent of Resection of Contrast-Enhanced and Non-Contrast-Enhanced Tumor With Survival Within Molecular Subgroups of Patients With Newly Diagnosed Glioblastoma. *JAMA Oncol.* 2020:10.1001/jamaoncol.2019.6143.
21. Jenkinson MD, Barone DG, Bryant A, et al. Intraoperative imaging technology to maximise extent of resection for glioma Review. *Cochrane Database Syst Rev.* Jan 22 2018;1:CD012788.
22. Sanai N, Mirzadeh Z, Berger MS. Functional outcome after language mapping for glioma resection. *N Engl J Med.* 2008;358(1):18-27.
23. Pessina F, Navarria P, Cozzi L, et al. Maximize surgical resection beyond contrast-enhancing boundaries in newly diagnosed glioblastoma multiforme: is it useful and safe? A single institution retrospective experience. *J Neuro-Oncol.* 2017;135(1):129-139.
24. Verburg N, Koopman T, Yaqub MM, et al. Improved detection of diffuse glioma infiltration with imaging combinations: a diagnostic accuracy study. *Neuro Oncol.* Sep 24 2019.
25. Kubben PL, Postma AA, Kessels AGH, Van Overbeeke JJ, Van Santbrink H. Intraobserver and interobserver agreement in volumetric assessment of glioblastoma multiforme resection. *Neurosurgery.* 2010;67(5):1329-1334.

Supplementary materials

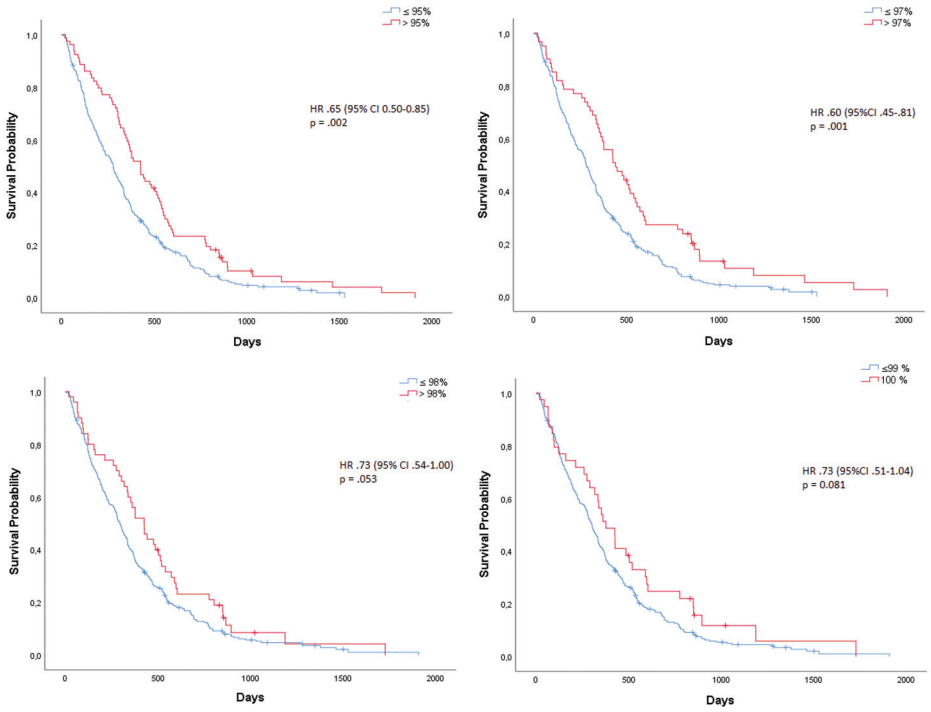


Figure S1. Kaplan Meier curves of OS for different EOR thresholds.

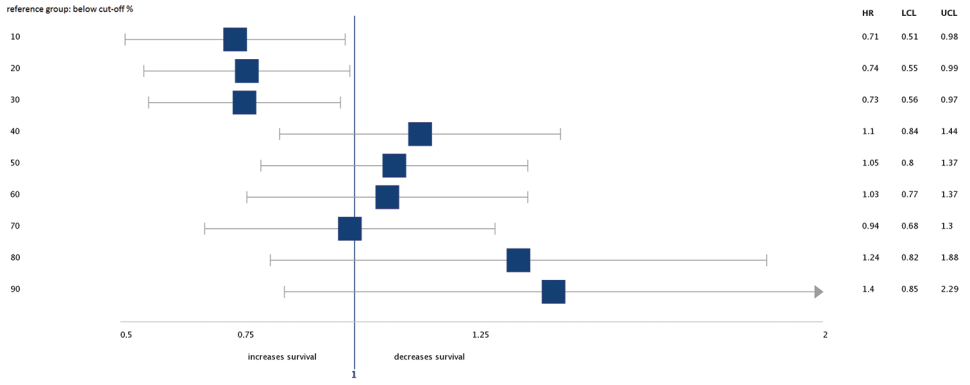


Figure S2. Hazard ratios with 95% CI for different NCE resection thresholds in explorative multivariable Cox proportional hazards models (stratified for MGMT methylation status, adjusted for age, KPS, adjuvant therapy, EOR and pre-operative tumor volume).

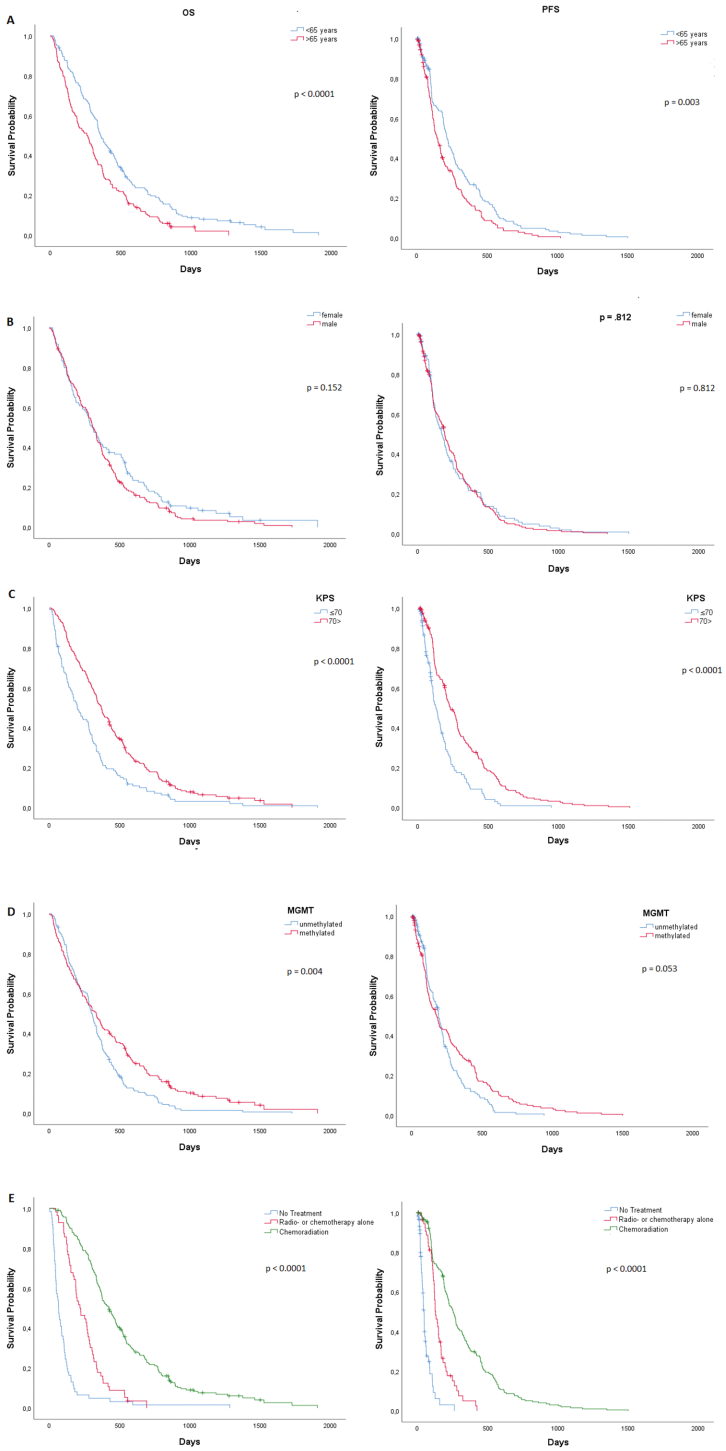


Figure S3. Kaplan Meier curves of OS and PFS for (A) age, (B) gender, (C) KPS, (D) MGMT promotor methylation, (E) adjuvant therapy.

Chapter 9

Development and external validation of a clinical prediction model for survival in glioblastoma patients

Hendrik-Jan Mijderwijk, Daan Nieboer, Fatih Incekara, Berger K, Daniel Hänggi, Michael S Sabel, Jörg Felsberg, Guido Reifenberger, Martin J van den Bent, Marion Smits, Marie-Therese Forster, Marcel A Kamp

Submitted

Abstract

Background

Glioblastoma survival prognostication has become more refined by the molecular reclassification into isocitrate dehydrogenase (IDH) wild-type and IDH mutant tumors. We aimed to provide an updated prediction model that predicts individual survival prognosis in IDH wild-type glioblastoma patients.

Methods

Data from existing databases from Germany and The Netherlands provided data on de novo diagnosed glioblastoma patients treated between 2012 and 2018. The prediction model considered recent glioblastoma biology markers in addition to well-known classical prognostic variables which were updated and refined with additional categories. The clinical prediction model was developed with Cox proportional hazards regression. Performance was evaluated according to calibration (calibration plots, calibration slope) and discrimination (c-statistic) in a cross-validation procedure to assess external validity.

Results

The German patient set consisted of 710 patients of whom 511 (72%) had died. Median follow-up was 11 months. The Dutch patient set consisted of 326 patients of whom 308 (94.5%) had died. Median follow-up was 10 months. Patient sets (n=1036) were combined to develop three models in order of increasing complexity. The final model considering age, gender, preoperative Karnofsky performance status, extent of surgical resection, MGMT promoter methylation status, and adjuvant therapeutic regime showed an optimism-corrected c-statistic of 0.73 (95% confidence interval 0.71 – 0.75). Cross-validation between the national cohorts yielded comparable results. Moderate miscalibration was observed.

Conclusion

The prediction model reliably predicts individual survival prognosis in newly diagnosed IDH wild-type glioblastoma patients, although additional validation for long-term survival may be desired. After further validation, the nomogram and web application (https://www.evidencio.com/models/show/2384) support shared decision making.

Introduction

Glioblastoma is the most common primary malignant brain tumor and the third most frequently reported central nervous system (CNS) tumor.[1] Its annual age-adjusted incidence rate of 3.21 per 100,000 person-years is the highest among malignant brain and CNS tumors patients.[1] Patients suffering from glioblastoma face a poor survival prognosis with a 5-year survival rate of less than 10%.[2] Nonetheless, interpatient variability in survival is substantial which is partly due to the tumor biology.[3]

In 2016, the revised 4th edition of the World Health Organization (WHO) classification of CNS tumors created two distinct glioblastoma entities according to isocitrate dehydrogenase (IDH) mutation status.[4] The vast majority of glioblastoma patients (>90%) harbor a IDH wild-type status and have a median overall survival equaling 1.2 years.[3] In contrast, glioblastoma patients molecularly labelled with a IDH mutation are less common (<10%) reaching a favorable 3-fold increased median overall survival time when compared to IDH wild-type glioblastoma patients.[3] Despite this categorisation, survival in IDH wild-type glioblastoma patients remains diverse. Prognostication of individual patient survival times depend on a range of prognostic variables related to patient characteristics, neurosurgical approach, glioblastoma biology, and adjuvant treatment strategies.[5-9]

Accurate prediction of individual patient survival is vital for personalised medicine and shared decision making. Nowadays, it is imperative to facilitate shared decision making, i.e. to inform patients and their relatives so that they understand their risk making conjoint decisions on choices possible. Consequently, therapeutic regimes can be better tailored to the individual patient and clinical scenario.

Clinical prediction models and their visualization, especially nomograms, are powerful tools for individualised estimation of patient survival times and patient counseling. However, since the recategorisation of glioblastoma, only a few reports address the use of nomograms for newly diagnosed glioblastoma patients.[9,10] Although the highest incidence rates of brain and CNS cancer are mainly documented in Europe,[11] a validated nomogram developed in European patients does currently not exist. We aimed to develop and externally validate a clinical prediction model to better predict survival in Western-European patients diagnosed with de novo glioblastoma without alterations in the IDH gene, considering traditional and modern predictors.

Methods

Study design and population

Glioblastoma patients from three university hospitals in Western-Europe were selected for model development and validation (University Medical Center Düsseldorf, and Frankfurt, Germany; and Erasmus MC, the Netherlands). Patients were eligible for analysis if they were at least 18 years of age at the day of neurosurgical intervention, and were

histopathologically diagnosed with de novo IDH wild-type glioblastomas according to the WHO classification of CNS tumors 2016 as recently recommended.[4,9] Glioblastomas were histopathologically classified as IDH wild-type, WHO grade IV according to the WHO classification of central nervous system tumors 2016.[4] Glioblastomas from patients diagnosed before 2016 were neuropathologically re-evaluated and reclassified according to the WHO 2016 criteria. The IDH mutation status was assessed by immunohistochemistry for IDH1-R132H as previously recommended.[12,13] Tumors of patients younger than 55 years of age were additionally investigated for less common mutations at codon 132 of IDH1 and codon 172 of IDH2 by Sanger sequencing or pyrosequencing.[12] Patients were excluded from analysis if a neurosurgical resection was performed more than 4 weeks after a biopsy procedure. The development set included patients from University Medical Center Düsseldorf (n=279) collected from 2013 – 2018 and from University Medical Center Frankfurt (n=431) collected from 2012 – 2018. The validation set was derived from Erasmus MC including 326 patients collected from 2012 – 2018.[14]

Ethics approval for the study was obtained from the institutional review boards at each center, i.e. the University Medical Center Düsseldorf (2019-474-RetroDEuA), University Medical Center Frankfurt (SNO-12-2019), and Erasmus MC (MEC-2019-0641).

Outcome definition

Overall survival was assessed from the day of first surgery until death or last follow-up. Patients were censored at the date of last follow-up.[7]

Candidate prognostic variables

Based on literature review and subject matter knowledge we considered the following predictor variables:

- Patient characteristics: gender, age, preoperative performance status (Karnofsky performance status (KPS)) were collected by reviewing patient charts. KPS was assessed preoperatively at the day of admission.
- Surgical results: extent of surgical resection (EOR). EOR was defined as gross total resection (GTR), non-GTR, and biopsy.[8] GTR was defined as complete removal of contrast enhancement on early T1-weighted postoperative MRI imaging (<72h) by a neuroradiologist blinded to intraoperative and histopathological findings.[8,15]
- Glioblastoma biology and adjuvant treatment strategies: O^6-methylguanine-DNA methyltransferase (MGMT) promoter methylation status, and adjuvant therapeutic regime. MGMT promoter methylation status was determined by pyrosequencing of sodium bisulfite-treated DNA and/or methylation specific PCR (MSP) as previously reported.[16,17] Adjuvant therapeutic regime was defined as Stupp, non-Stupp, and none.[8] The Stupp category consisted of radiotherapy plus concomitant and maintenance temozolomide.[5] The non-Stupp category consisted of subparts/ modifications of the Stupp protocol and experimental designs.[8] Decisions on therapy

were rendered by the local multidisciplinary tumor boards and analysed according to an intention-to-treat principle.

Sample size

Conventional sample size recommendations require at least 10 to 20 events per candidate prognostic variable, which target was easily met.[18] In addition, we performed a more advanced calculation.[19] Using the observed *c*-statistic from Gittleman et al.[9] (*c*-statistic 0.76 with 163 events) we would need more than 200 patients to ensure a heuristic shrinkage slope above 0.9 for the prediction model. For models with a lower *c*-statistic at least 300-500 patients would be required for reliable modeling.

Statistical analysis

For continuous data we used means, standard deviations and ranges. For categorical data we used counts and percentages. The clinical prediction model generation was in accordance to recent methodology,[18,20] with reporting according to the TRIPOD guidelines.[21,22]

Model development

Cox regression was used to develop a clinical prediction model estimating survival. Age and preoperative KPS were kept as continuous prognostic variable in the analysis to avoid loss of prognostic information.[23] In addition, we explored non-linearity for the association of age and preoperative KPS with mortality using restricted cubic splines.[24] Missing values were assumed to be missing at random and multiple imputation was performed using the mice algorithm.[25] Missing values were imputed 10 times. Statistical analyses were performed on the 10 imputed dataset and results were pooled using Rubin rules. The modeling procedure consisted of three models of increasing complexity:

1. Clinical model: including age, gender, and KPS.
2. Surgical model: adding EOR to the clinical model.
3. Treatment model: consisting of the surgical model plus MGMT methylation status and adjuvant treatment regime.

Hazard ratios (HRs) with 95% confidence intervals (95% CI) were estimated as measures for association of the prognostic variables with survival.

Model performance

We assessed the quality of the prediction model by evaluating calibration and discrimination measures.[18] Model calibration gauges the agreement between the predictions of the model with the observed survival probability.[18] Model calibration was graphically assessed by calibration plots.[26] Differences in baseline risk were studied by adding cohort as a factor in the model. Furthermore, the calibration slope was calculated.[27]

Harrell's concordance statistic (*c*-statistic) was used to quantify the discriminative ability.[28] Model discrimination tells us how good the constructed prediction model identifies—for two randomly chosen patients—the patient who deceased first with a higher probability of dying. An uninformative model will have a *c*-statistic of 0.5, whereas a model with perfect discrimination will have a *c*-statistic equaling 1.0.[18] Furthermore, we quantified the heterogeneity in case-mix in the development and validation population to aid interpretation of the observed *c*-statistic at validation.[29]

Model validation

We first developed models in Germany with validation in Erasmus MC. We then reversed this procedure with development in Erasmus MC and validation in Germany. This cross-validation procedure between the national datasets shows the external validity of the prediction model.[30] Subsequently, the final model was developed on the combined data, provided that no major between cohort heterogeneity was found. The performance of this final model was estimated by a bootstrap (1000 samples) validation procedure.[30]

Model presentation

Nomograms were created to predict an individual patient's median, 1-year, and 2-year survival probabilities. Descriptive analysis and prediction modeling analysis were performed using R software version 3.5.2. The significance level was set at 5% for all analyses.

Results

Study population

The combined German data set consisted of data from 713 patients. We excluded 3 patients in whom a surgical resection was performed 4 weeks after the initial biopsy. Thus, after imputation, 710 complete cases were analysed. The Erasmus MC data set did not have any missing value and consisted of 326 patients. The overall survival as assessed by the Kaplan-Meier method is visualized in figure 1. In both data sets, age at diagnosis was comparable and more than half of the patients was male (table 1). In the German data set, 85% of the patients had a preoperative KPS greater than 70, whereas in the validation set 64% of the patients had a preoperative KPS above 70. A minority of patients had complete resection (ca. one fifth in the development set and one tenth in the validation set), whereas the majority of patients had partial resection followed by a biopsy-only procedure. Approximately half of the patients were labelled with a MGMT promotor-methylated tumor. The majority of the patients received the Stupp regime as post-surgical therapy. However, in the development set more patients were assigned to a non-Stupp regime (32% vs. 17%). Nearly one-fifth of the patients were not assigned to postoperative therapy.

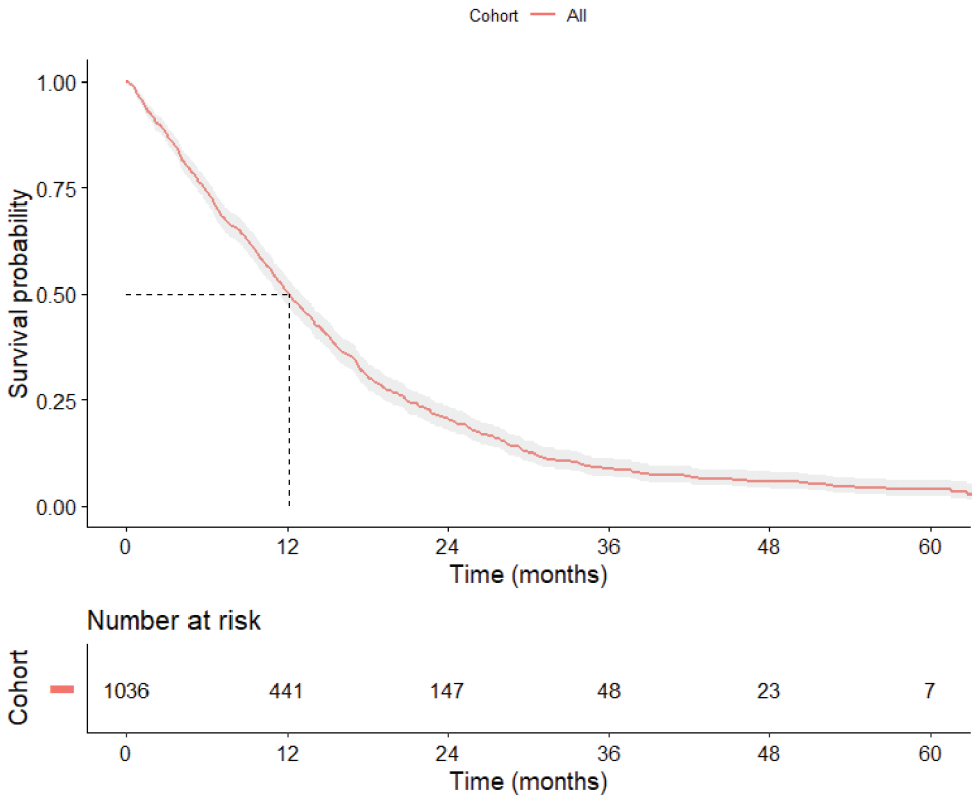


Figure 1. Kaplan-Meier estimate of survival.

The median duration of follow-up of survivors in years was nearly the same in both sets, 0.89 vs. 0.84 (table 1). 308 patients had died in the cohort from the Netherlands compared to 511 of the German cohorts. Univariable hazard ratios (HRs) between the predictors and mortality can be found in table 2. Table S1 shows more details on patients' characteristics stratified according to the respective academic center.

Model development, performance, validation, and presentation

Age and preoperative KPS could be modelled well assuming a linear association. Besides patient gender (HR = 0.94 (0.79 – 1.11)), all prognostic variables showed statistically significant associations with survival in the developed model (figure 2). Younger patients at diagnosis with a higher preoperative KPS had better survival (HR age per year = 1.32 (1.22 – 1.42); HR preoperative KPS = 0.85 (0.76 – 0.94)). Incomplete surgical tumor resection (HR partial resection = 1.30 (1.04 – 1.64); HR biopsy-only = 1.95 (1.52 – 2.49)) and deviations from standard adjuvant therapy (HR non-Stupp = 1.29 (1.06 – 1.58); HR no therapy = 2.38 (1.85 – 3.07)) were statistically significantly associated with worse survival. Patients labelled with a MGMT promotor-methylated tumor confer a favorable survival

prognosis compared to those labelled with a MGMT promotor-unmethylated tumor (HR 0.50 (0.41 – 0.62), table S2).

The direction of the predictor effects was the same in the German development set and the Dutch validation set (figure 2, table S3). The treatment model had somewhat stronger effects in the Dutch cohort (interaction by cohort: p<0.001) but not for the clinical and surgical model (p=0.068 and p=0.248) without any obvious reason.

Table 1. Patient characteristics of the data used at model development.

	Germany (n=710)	**Netherlands (n=326)**	**P-value**
Time of follow-up of survivors (median [IQR])	0.89 [0.37, 1.49]	0.84 [0.37, 1.43]	<0.001
Died (%)	511 (72.0)	308 (94.5)	<0.001
Age (median [IQR])	64 [55, 73]	65 [57, 72]	0.864
Male sex (%)	385 (54.2)	206 (63.2)	0.008
Preoperative KPS (%)			<0.001
70 or lower	88 (15.1)	119 (36.5)	
80	93 (16.0)	92 (28.2)	
90	218 (37.4)	85 (26.1)	
100	184 (31.6)	30 (9.2)	
Surgical resection (%)			<0.001
Complete resection	158 (22.4)	34 (10.4)	
Partial resection	321 (45.5)	214 (65.6)	
Biopsy-only	227(32.2)	78 (23.9)	
MGMT methylated (%)	291 (47.2)	177 (54.3)	0.044
Adjuvant therapy (%)			<0.001
No therapy	117 (16.8)	61 (18.7)	
Non-Stupp	220 (31.6)	56 (17.2)	
Stupp	360 (51.6)	209 (64.1)	

P-values for comparison were calculated with the t test, Mann-Whitney *U* test, or χ^2 test. IQR=interquartile range. KPS= Karnofsky performance status. MGMT= O^6-methylguanine-DNA methyltransferase.

Table 2. Univariable association between predictors and mortality (n=1036).

Predictor	Measure/category	HR (95% CI)
Age	Per decade	1.33 (1.25, 1.41)
Sex	Male vs. female	1.10 (0.96, 1.26)
Preoperative KPS	Per 10 points increase	0.75 (0.70, 0.80)
Surgical resection	Complete	reference
	Partial	1.50 (1.24, 1.81)
	Biopsy-only	2.38 (1.94, 2.92)
MGMT	Methylated vs. wild-type	0.65 (0.56, 0.75)
Adjuvant therapy	Stupp	reference
	Non-Stupp	1.48 (1.26, 1.73)
	No therapy	3.90 (3.23, 4.71)

HR=hazard ratio. CI=confidence interval. KPS= Karnofsky performance status. MGMT= O^6-methylguanine-DNA methyltransferase.

The apparent *c*-statistic of the developed prediction models in the development set was promising with the treatment model having the highest discriminative ability (*c*-statistic 0.74, 95% CI 0.71 – 0.76, table S2).

Validation

At cross-validation by county, we confirmed an increasing *c*-statistic with increasing model complexity (table 4). The *c*-statistic for the treatment model was 0.70 (95% CI 0.67 – 0.73) indicating a well discriminating prediction model. The calibration plots showed some miscalibration, especially for predicting long-term survival probabilities (figure 3). The calibration plots suggested that the clinical model underestimated survival while the surgical and treatment model overestimated survival (figure 3). In more detail, beyond one year of survival follow-up the predicted curve deviated more from the observed curve. The calibration slope was around 1 for all models: <1 for the clinical and surgical model and >1 for the treatment model (table 4). The apparent *c*-statistics when refitting the models in the Erasmus MC data were 0.62, 0.64, and 0.73 respectively (table S2). When we reversed the validation procedure, the validated *c*-statistics in the Germany data set were 0.65 (0.62 – 0.67), 0.68 (0.66 – 0.71), and 0.72 (0.69 – 0.74) for the clinical, surgical, and treatment models respectively (table 4). The calibration plot for the clinical model showed an excellent agreement between observed and survival probability (figure 3).

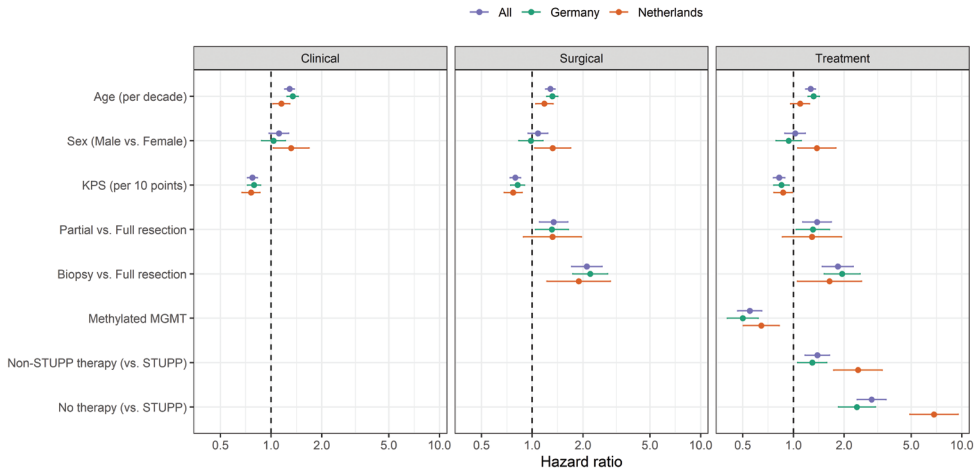


Figure 2. Association of the prognostic variables with survival.

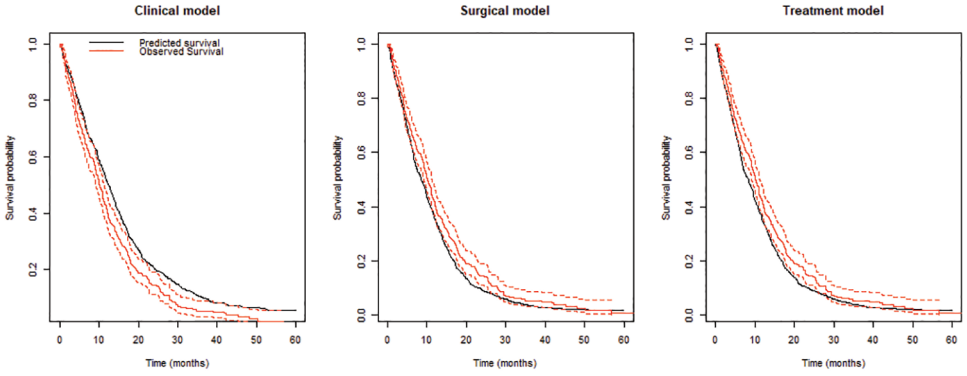


Figure 3. Calibration plots of the developed models at cross-validation.

Table 3. Hazard ratio and associated 95% confidence intervals in the final prediction models (n=1036) including the discriminative ability.

Predictor	Measure/Category	Model		
		Clinical	**Surgical**	**Treatment**
Age	Per decade	1.29 (1.21, 1.37)	1.28 (1.20, 1.37)	1.27 (1.19, 1.35)
Sex	Male vs. female	1.11 (0.97, 1.27)	1.08 (0.94, 1.24)	1.02 (0.89, 1.18)
Preoperative KPS	Per 10 points increase	0.78 (0.72, 0.83)	0.79 (0.74, 0.85)	0.82 (0.76, 0.89)
Surgical resection	Complete		reference	reference
	Partial		1.34 (1.11, 1.62)	1.38 (1.14, 1.68)
	Biopsy		2.11 (1.71, 2.59)	1.84 (1.48, 2.27)
MGMT	Wild-type vs methylated			0.55 (0.47, 0.65)
Adjuvant therapy	Stupp			reference
	Non-Stupp			1.39 (1.18, 1.64)
	No therapy			2.92 (2.39, 3.55)
c-statistic		0.66 (0.64, 0.68)	0.68 (0.66, 0.70)	0.73 (0.71, 0.75)

KPS= Karnofsky performance status. MGMT= O^6-methylguanine-DNA methyltransferase.

Again, the surgical and treatment model overestimated survival yet more pronounced, especially for predicting long-term survival (figure 3).

The spread between the predictions (standard deviation of the linear predictor) increased with model complexity, and was less in the validation set for all models (table 5). This indicates that the decrease in *c*-statistic was partly due to a decrease in case-mix heterogeneity form development to validation set.

Model presentation

The final model combined all data from the development set and validation set yielding comparable associations of the prognostic variables with survival (table 3, figure 2). The *c*-statistic equaled 0.73 (0.71 – 0.75). We developed nomograms to predict a patient's individual survival rate for several time periods (figures 4-6). In addition, an online prognostic calculator based on the model algorithms including error margins (95% CI for prediction) is accessible at https://www.evidencio.com/models/show/2384 and shown in the Box. Table S3 provides the baseline hazard and predictor coefficients for the different models to allow for independent external validation studies by independent researchers.

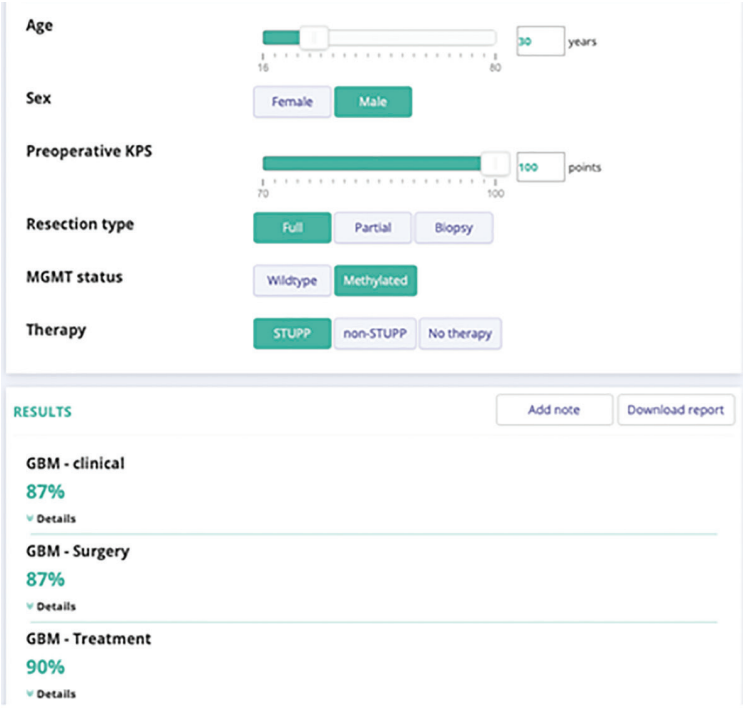
Table 4. Performance of the developed prediction models at external validation in the Dutch dataset.

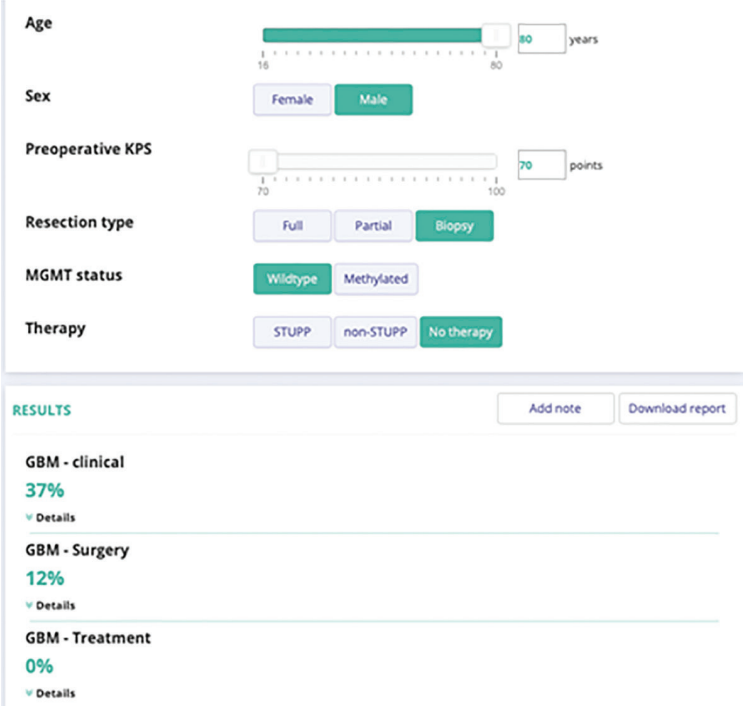
Validation set	Performance measure		Model
	calibration slope (95% CI)	c-statistic (95% CI)	
Netherlands	0.73 (0.43, 1.03)	0.61 (0.58, 0.65)	Clinical
Germany	1.02 (0.78, 1.26)	0.65 (0.62, 0.67)	Clinical
Netherlands	0.81 (0.52, 1.09)	0.62 (0.59, 0.66)	Surgical
Germany	1.01 (0.82, 1.20)	0.68 (0.66, 0.71)	Surgical
Netherlands	1.12 (0.87, 1.36)	0.70 (0.67, 0.73)	Treatment
Germany	0.67 (0.57, 0.76)	0.72 (0.69, 0.74)	Treatment

Table 5. Standard deviation of the linear predictor for the different models in the development and validation set.

	Germany (development)	The Netherlands (validation)
Clinical model	0.45	0.39
Surgical model	0.57	0.45
Treatment model	0.78	0.71

Box. Example patients for good and poor one-year survival probability.





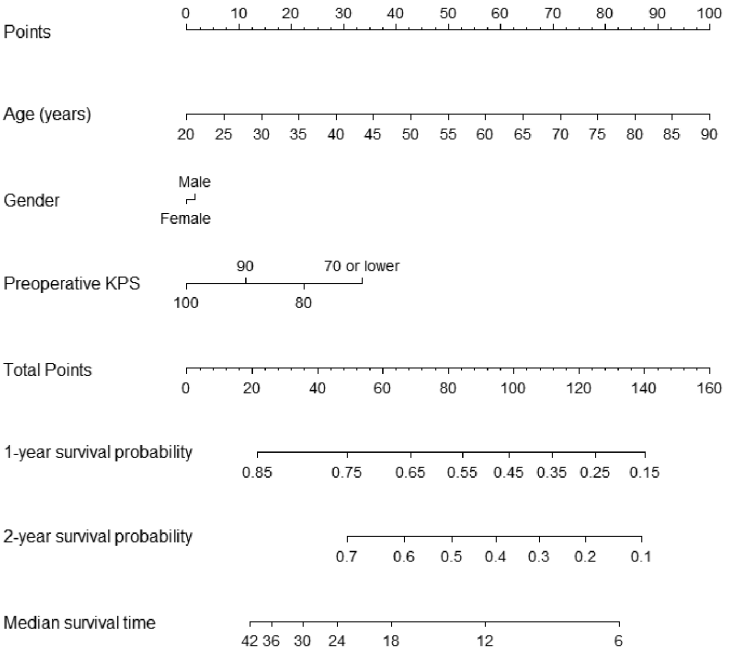


Figure 4. Nomogram clinical model.

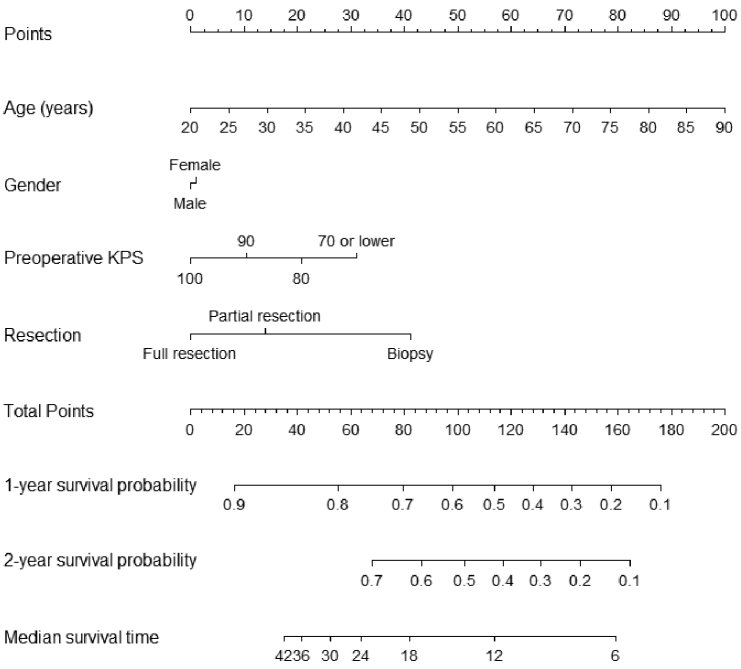


Figure 5. Nomogram surgical model.

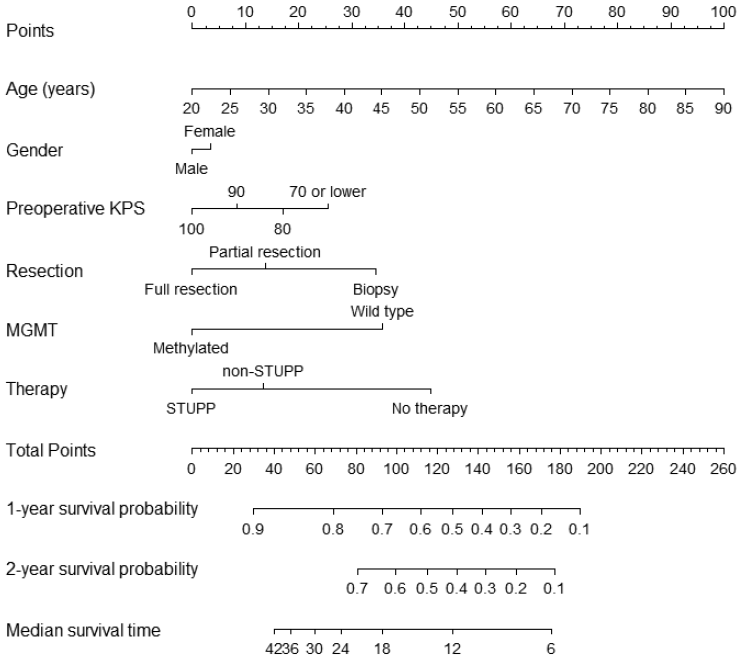


Figure 6. Nomogram treatment model.

Discussion

Individualised estimates of survival time can be obtained with reasonable accuracy from the proposed clinical prediction model in patients with newly diagnosed with IDH wild-type glioblastoma. The model is cross-validated in patient cohorts from Germany and The Netherlands. Therefore, the model may be particularly useful for medical- and shared decision making in Western-European patients. The web based prognostic calculator will facilitate clinical uptake. Considering updated conventional predictors and new predictor variables including current glioblastoma biology, the prediction model reached a promising discriminative model performance (*c*-statistic 0.73).

Before publication of the revised 4th edition of the WHO classification of CNS tumors in 2016, most literature reporting on prediction models for glioblastoma patients is confounded by omitting information on glioblastoma biology in the analyses.[7,31] Recently, Gittleman et al. developed and validated a clinical prediction model in a Northern American population that does take into account glioblastoma biology.[9] We found similar predictor effects for age at diagnosis, patient gender, and preoperative KPS, although we avoided to dichotomize preoperative KPS to prevent extensive loss of data. About half of the patients had a MGMT promoter methylation status, in keeping with previous evidence.[6] The predictor effect of MGMT promotor methylation was also in line with the work by Gittleman et al.[9] Furthermore, as recently recommended, we expanded and updated the model with an additional surgical intervention, i.e. the biopsy-only group, and an additional adjuvant treatment option, i.e. the non-Stupp alternative. As expected, patients undergoing the biopsy-only variant had a worse prognosis. Patient allocated to the non-Stupp alternative treatment confer a favorable survival prognosis compared to those patients having no adjuvant therapy. To address the addition of multiple parameters to the model, a more robust effective sample size was achieved to provide accurate predictions. Nonetheless, the work by Gittleman et al.[9] presented a higher *c*-statistic of 0.75 at external validation. This could be due to a lower case-mix heterogeneity in the present validation set. The lower *c*-statistic may also be explained by measurement error that might have been emerged in the present study.

The assessment of *extent of surgical resection* differed between the model development set and the model validation set. In the former, extent of surgical resection was defined by a qualitative approach,[15] whereas in the latter a quantitative (volumetric segmentation analysis) was used.[14] Consequently, the measurement error in the validation set is likely lower possibly resulting in a different association between *extent of surgical resection* and mortality. Furthermore, we found a stronger effect—yet puzzling result—of the adjuvant treatment on patient survival in the validation set compared to the patients from the model development set. Local patient allocation to the adjuvant treatment groups according to local principles and the miscellaneous option/non-STUPP might have induced heterogeneity between the datasets in the adjuvant treatment variable.

The number of patients that did not complete the Stupp protocol are unknown. Although speculating, it might be that more patients in the development set were non-compliant when compared to patients in the validation set, which might have resulted in a different association with mortality. Other unknown factors not captured in this study might have affected patient allocation to adjuvant treatment.

Implications for patient management

Patients newly diagnosed with IDH wild-type glioblastoma need to be well informed about the prognosis of this devastating disease. To participate adequately in shared decision making, patients and their relatives need to understand their prognosis to make preference-sensitive decisions. Since the new WHO categorisation, an updated prediction tool is inevitable for providing reliable predictions. The proposed prediction model is particularly useful for shared decision making. The nomograms and online calculator presented here are intuitive and freely available to facilitate shared decision making in the clinical setting. The different models can be used preoperatively and postoperatively by health care professionals to explain the clinical scenario expected. Consequently, patient-tailored treatment guidance and future planning becomes better feasible.

Implications for future work

Nowadays, glioblastoma cannot be treated as a signal histopathological entity. Although the presented model addresses glioblastoma as a molecular heterogeneous entity, future model updating is likely necessary, e.g. taking into account newly defined molecular subgroups of IDH wild-type glioblastoma characterized by distinct DNA methylome profiles, or other potential biomarkers like tumor mutational burden or total copy number aberration.[37] Along with this basic scientific research, ensuing clinical therapies are designed and tested. Tumor treating fields concurrent with temozolomide have been suggested to be effective.[33] Immunotherapies and precision oncology approaches have so far not shown to increase survival.[32] If those therapies become standard care, model updating will likely further increase the predictive performance.

Strengths and limitations

A strength of the current study is the development and validation in geographically distinct settings. Other centers may have different case-mix and different local care pathways, which may threaten external validity. The generalizability of the model to non-academic centers needs to be tested in future work. As the model was generated in Western-Europe, application of the model in other geographical areas should be performed with caution. Second, although the present model updated some conventional prognostic variables and did take into account relevant molecular biomarkers, the model performance was not perfect. Other predictors may need to be considered such as corticosteroid use, seizures and hospital complications including venous tromboembolism.[34-36] These events

may be associated with outcome and hence affect the accuracy of the presented models. However, including such events may make a clinical prediction very difficult to apply since these data is generally not available at baseline. Third, we cannot rule out some information bias, since some variables have been collected retrospectively. Fourth, by lack of a comparative design in the datasets, a causal relationship between treatment and prediction cannot be shown. Therefore, this work is especially useful for shared decision making and has the potential to be a basis for impact studies on personalised medicine purposes. Fifth, the Erasmus MC cohort did not have any missing data, and the German data was nearly complete for all cases. Multiple imputation is advisable to prevent loss of prognostic information.[20] Nevertheless, some level of inaccuracy of the imputed data cannot be ruled out.

Finally, the model was developed and validated within a large sample size. However, some model miscalibration emerged, especially beyond the first year of survival. This may be due to a drop in the sample size as a substantial part of the patients has been deceased according to median survival times <14 months.

Conclusion

The proposed clinical prediction model reliably predicts individual survival prognosis in newly diagnosed Western-Europe IDH wild-type glioblastoma patients. The model paves the way for patient-tailored precision medicine and provides a framework that can be used for future updating. For clinical uptake, free software is available to facilitate medical and shared decision making at https://www.evidencio.com/models/show/2384, although additional validation may be desired, especially for predicting long-term survival.

References

1. Ostrom QT, Gittleman H, Truitt G, Boscia A, Kruchko C, Barnholtz-Sloan JS. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011-2015. *Neuro-Oncology* 2018; **20**: iv1–iv86.
2. Stupp R, Hegi ME, Mason WP, *et al.* Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol* 2009; **10**: 459–66.
3. Molinaro AM, Taylor JW, Wiencke JK, Wrensch MR. Genetic and molecular epidemiology of adult diffuse glioma. *Nat Rev Neurol* 2019; **15**: 405–17.
4. 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**: 803–20.
5. Stupp R, Mason WP, van den Bent MJ, *et al.* Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med* 2005; **352**: 987–96.
6. Hegi ME, Diserens A-C, Gorlia T, *et al.* MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 2005; **352**: 997–1003.
7. Gorlia T, van den Bent MJ, Hegi ME, *et al.* Nomograms for predicting survival of patients with newly diagnosed glioblastoma: prognostic factor analysis of EORTC and NCIC trial 26981-22981/CE.3. *Lancet Oncol* 2008; **9**: 29–38.
8. Gessler F, Bernstock JD, Braczynski A, *et al.* Surgery for Glioblastoma in Light of Molecular Markers: Impact of Resection and MGMT Promoter Methylation in Newly Diagnosed IDH-1 Wild-Type Glioblastomas. *Neurosurgery* 2019; **84**: 190–7.
9. Gittleman H, Cioffi G, Chunduru P, *et al.* An independently validated nomogram for isocitrate dehydrogenase-wild-type glioblastoma patient survival. *Neuro-Oncology Advances* 2019; **1**: 1985–9.
10. Shen E, Johnson MO, Lee JW, *et al.* Performance of a nomogram for IDH-wild-type glioblastoma patient survival in an elderly cohort. *Neuro-Oncology Advances* 2019; **1**: 998–2.
11. Miranda-Filho A, Piñeros M, Soerjomataram I, Deltour I, Bray F. Cancers of the brain and CNS: global patterns and trends in incidence. *Neuro-Oncology* 2016; **50**: now166–11.
12. Felsberg J, Wolter M, Seul H, *et al.* Rapid and sensitive assessment of the IDH1 and IDH2 mutation status in cerebral gliomas based on DNA pyrosequencing. *Acta Neuropathol* 2010; **119**: 501–7.
13. Hartmann C, Hentschel B, Wick W, *et al.* Patients with IDH1 wild type anaplastic astrocytomas exhibit worse prognosis than IDH1-mutated glioblastomas, and IDH1 mutation status accounts for the unfavorable prognostic effect of higher age: implications for classification of gliomas. *Acta Neuropathol* 2010; **120**: 707–18.
14. Incekara F, Smits M, van der Voort SR, *et al.* The Association Between the Extent of Glioblastoma Resection and Survival in Light of MGMT Promoter Methylation in 326 Patients With Newly Diagnosed IDH-Wildtype Glioblastoma. *Front Oncol* 2020; **10**: 987–8.

15. Senft C, Bink A, Franz K, Vatter H, Gasser T, Seifert V. Intraoperative MRI guidance and extent of resection in glioma surgery: a randomised, controlled trial. *Lancet Oncol* 2011; **12**: 997–1003.
16. Felsberg J, Rapp M, Loeser S, *et al.* Prognostic significance of molecular markers and extent of resection in primary glioblastoma patients. *Clinical Cancer Research* 2009; **15**: 6683–93.
17. Felsberg J, Thon N, Eigenbrod S, *et al.* Promoter methylation and expression of MGMT and the DNA mismatch repair genes MLH1, MSH2, MSH6 and PMS2 in paired primary and recurrent glioblastomas. *Int J Cancer* 2011; **129**: 659–70.
18. Mijderwijk H-J, Steyerberg EW, Steiger H-J, Fischer I, Kamp MA. Fundamentals of Clinical Prediction Modeling for the Neurosurgeon. *Neurosurgery* 2019; **85**: 302–11.
19. Riley RD, Ensor J, Snell KIE, *et al.* Calculating the sample size required for developing a clinical prediction model. *BMJ* 2020; **368**: m441–12.
20. Steyerberg EW. Clinical Prediction Models. Cham: Springer International Publishing, 2019 DOI:10.1007/978-3-030-16399-0.
21. Collins GS, Reitsma JB, Altman DG, Moons K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med* 2015; **13**: 1–10.
22. Heus P, Reitsma JB, Collins GS, *et al.* Transparent Reporting of Multivariable Prediction Models in Journal and Conference Abstracts: TRIPOD for Abstracts. *Ann Intern Med* 2020; **173**: 42–7.
23. Collins GS, Ogundimu EO, Cook JA, Manach YL, Altman DG. Quantifying the impact of different approaches for handling continuous predictors on the performance of a prognostic model. *Stat Med* 2016; **35**: 4124–35.
24. Harrell FE, Lee KL, Pollock BG. Regression models in clinical studies: determining relationships between predictors and response. *JNCI Journal of the National Cancer Institute* 1988; **80**: 1198–202.
25. Van Buuren S. Flexible imputation of missing data. 2018.
26. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol* 2013; **13**: 33–15.

Supplementary materials

Table S1. Patient characteristics stratified per hospital.

	Düsseldorf (n=279)	Frankfurt (n=431)	Erasmus MC (n=326)
Time of follow-up of survivors (median [IQR])	0.96 [0.49, 1.57]	0.82 [0.33, 1.45]	0.84 [0.37, 1.43]
Died (%)	192 (68.8)	319 (74.0)	308 (94.5)
Age (median [IQR])	65.00 [55, 73]	64 [55, 72]	65 [57, 72]
Male sex (%)	115 (41.2)	270 (62.6)	206 (63.2)
Preoperative KPS (%)			
70 or lower	27 (9.7)	61 (20.0)	119 (36.5)
80	60 (21.6)	33 (10.8)	92 (28.2)
90	110 (39.6)	108 (35.4)	85 (26.1)
100	81 (29.1)	103 (33.8)	30 (9.2)
Surgical resection (%)			
Complete resection	91 (32.9)	67 (15.6)	34 (10.4)
Partial resection	170 (61.4)	151 (35.2)	214 (65.6)
Biopsy-only	16 (5.8)	211 (49.2)	78 (23.9)
MGMT methylated (%)	129 (46.7)	162 (47.5)	177 (54.3)
Adjuvant therapy (%)			
No therapy	38 (13.6)	79 (18.9)	61 (18.7)
Non-Stupp	83 (29.7)	137 (32.8)	56 (17.2)
Stupp	158 (56.6)	202 (48.3)	209 (64.1)

IQR=interquartile range. KPS= Karnofsky performance status. MGMT= O^6-methylguanine-DNA methyltransferase.

Table S2. Hazard ratio and associated 95% confidence intervals in the developed prediction models including the apparent discriminative ability.

Predictor	Measure/ Category	Germany *development set*			The Netherlands *validation set*		
		Model			Model		
		Clinical	Surgical	Treatment	Clinical	Surgical	Treatment
Age	Per decade	1.34 (1.25, 1.45)	1.32 (1.22, 1.42)	1.32 (1.22, 1.42)	1.15 (1.03, 1.29)	1.18 (1.05, 1.33)	1.10 (0.96, 1.25)
Sex	Male vs. female	1.03 (0.88, 1.22)	0.98 (0.83, 1.16)	0.94 (0.79, 1.11)	1.32 (1.03, 1.68)	1.32 (1.03, 1.69)	1.38 (1.06, 1.79)
Preoperative KPS	Per 10 points increase	0.79 (0.72, 0.87)	0.82 (0.74, 0.90)	0.85 (0.76, 0.94)	0.76 (0.67, 0.86)	0.77 (0.68, 0.87)	0.87 (0.77, 0.99)
Surgical resection	Complete		reference	reference		reference	reference
	Partial		1.31 (1.04, 1.64)	1.30 (1.04, 1.64)		1.32 (0.89, 1.96)	1.29 (0.86, 1.93)
	Biopsy		2.21 (1.74, 2.80)	1.95 (1.52, 2.49)		1.89 (1.22, 2.91)	1.64 (1.06, 2.54)
MGMT methylation	No vs yes			0.50 (0.41, 0.62)			0.64 (0.51, 0.82)
Adjuvant therapy	Stupp			reference			reference
	Non-Stupp			1.29 (1.06, 1.58)			2.42 (1.74, 3.37)
	No therapy			2.38 (1.85, 3.07)			6.83 (4.91, 9.49)
c-statistic		0.67 (0.64, 0.69)	0.70 (0.68, 0.72)	0.74 (0.71, 0.76)	0.62 (0.58, 0.65)	0.64 (0.60, 0.67)	0.73 (0.70, 0.76)

IQR=interquartile range. KPS= Karnofsky performance status. MGMT= O6-methylguanine-DNA methyltransferase.

Table S3. Coefficients and baseline hazard of the final updated prediction models.

Predictor	Measure/Category	Model		
		Clinical	Surgical	Treatment
Age	Per decade	0.25	0.25	0.24
Sex	Male vs. female	0.11	0.08	0.02
Preoperative KPS	Per 10 points increase	-0.25	-0.23	-0.19
Surgical resection	Complete		0	0
	Partial		0.29	0.32
	Biopsy-only		0.75	0.61
MGMT	Wild-type vs. methylated			-0.60
Adjuvant therapy	Stupp			0
	Non-Stupp			0.33
	No therapy			1.07
Baseline hazard at 1 year		0.50	0.49	0.48

IQR=interquartile range. KPS= Karnofsky performance status. MGMT= O^6-methylguanine-DNA methyltransferase.

Chapter 10.1

Association between supratotal glioblastoma resection and patient survival: a systematic review and meta-analysis

Fatih Incekara, Stephan Koene, Arnaud J.P.E. Vincent,
Martin van den Bent, Marion Smits

World Neurosurg. 2019 Jul;127:617-624

Abstract

Background

Gross total resection (GTR) of contrast enhancement (CE) improves survival of glioblastoma (GBM) patients. However, GBM infiltrates into brain parenchyma, beyond CE. It remains unclear whether resection beyond CE (supratotal resection, SPTR) improves survival without causing additional neurological deficits. The aim of this meta-analysis was to study the association between SPTR and overall survival in GBM patients.

Methods

Embase, Pubmed and other literature databases were searched for eligible studies until August 2018. Studies involving patients with GBM comparing SPTR with GTR were included in this study. The main outcome was overall survival as hazard ratios (HR) with 95% confidence interval (CI) and median overall survival differences with 95% CI.

Results

Meta-analysis based on six studies and 1,168 unique patients with GBM showed that when compared with GTR, SPTR of GBM resulted in 53% lower risk of mortality at any time during follow-up (HR = 0.47, 95% CI = 0.31 - 0.72, P = 0.0005). The median overall survival of SPTR was 6.4 months (95% CI = 3.2 - 9.7) longer than GTR (P = 0.0001). Reports on post-operative deficits were limited and the quality of evidence was moderate to very low.

Conclusion

When compared with GTR, SPTR of GBM suggests lower risk of mortality and longer median overall survival. However, the quality of evidence of the available studies was poor. Therefore, it remains unclear whether SPTR is safe and whether it really improves survival of patients with GBM. Future prospective trials and a standardized definition of SPTR are needed.

Introduction

Patients with glioblastoma (GBM) have a poor prognosis with a median overall survival of 15 months, despite safe and maximal surgical resection followed by chemo- and/or radiotherapy.[1] GBM typically appears on a post-contrast MRI scan as a contrast enhancing (CE) tumor with central necrosis. Maximal and safe surgical resection or gross total resection (GTR) of the CE area is currently the main goal of GBM surgery. Both the survival benefit of GTR and the optimal way to achieve this are extensively debated in the neurosurgical literature.

Brown et al. showed in a systematic review and meta-analysis of 37 articles and 41,117 unique patients, that in comparison with subtotal resection, GTR of CE area decreases the risk of mortality with 28% at 1 year and 16% at 2 years ($P < 0.001$).[1] Jenkinson et al. recently showed in a Cochrane review that resection with 5-Aminolevulinic Acid (5-ALA) fluorescence guidance or intra-operative MRI guidance may increase the extent of GBM resection.[2]

However, GBM is known to infiltrate far beyond the margins of CE as seen on MRI, into the surrounding edematous T2-FLAIR hyperintense region.[3] This raises the question whether GTR of the CE portion is indeed a 'total' resection, or whether surgical resection should also include –part of the- hyperintense T2-FLAIR region to improve survival. It is suggested that 5-ALA fluorescence accumulates in cancer cells and not only corresponds with the CE portion on MRI but also exceeds this area as vague fluorescence, which corresponds with portions of GBM often infiltrating into eloquent brain area.[4]

This concept of so-called supramarginal or supratotal resection (SPTR), is already known in the field of low grade glioma surgery,[5,6] because low grade gliomas commonly don't enhance and therefore lack the CE target for surgical resection. However, SPTR of GBM is less extensively investigated, and up until this date no quantitative data analysis has been performed to clarify the association of survival with SPTR versus GTR in GBM patients, nor of associated post-surgical neurological complications.

The aim of this study was to investigate the association between SPTR and survival in glioblastoma patients, using a systematic review and meta-analysis in accordance with the PRISMA guidelines.[7]

Methods

Systematic review

We searched for studies that included GBM patients who had received surgical resection, with pre- and post-operative MRI imaging. Letters, editorials, abstracts and non-English citations were excluded. The search query was designed together with an expert librarian at the Erasmus University Medical Center, Medical Library to capture all citations published until August 2018 within Pubmed [*(Glioblastoma/ OR (glioblastom* OR (maligna* ADJ3*

glioma) OR (high* ADJ3 grade* ADJ3 glioma*) OR ((grade-iv OR grade-4) ADJ3 glioma*) OR gbm).ab,ti,kw.) AND (Surgical Procedures, Operative/ OR exp Neurosurgery/ OR Neurosurgical Procedures/ OR exp Brain Neoplasms/su OR (surg* OR neurosurg* OR resect*).ab,ti,kw.) AND (Margins of Excision/ OR (flair OR (Fluid ADJ3 attenuat* ADJ3 invers* ADJ3 recover*) OR t2 OR t-2 OR gross-total OR ((exten* OR Supratotal* OR Supramaxim*) ADJ3 (resect* OR remov*)) OR ((surg* OR excis* OR resect*) ADJ3 margin*) OR ((beyond OR additional*) ADJ6 (contrast OR boundar*))).ab,ti,kw.) NOT (letter* OR news OR comment* OR editorial* OR congres* OR abstract* OR book* OR chapter* OR dissertation abstract*).pt.*] Embase, Web of Science, Cochrane Central and Google Scholar (search queries for these sources can be found in the Supplementary Materials).

After removing duplicate articles, two independent reviewers (F.I. and S.K.) screened the articles based on title and abstract, removing off-topic citations. Full texts of remaining articles were independently read by the reviewers to determine whether they were eligible for final inclusion. Articles that studied solely pediatric or non-glioblastoma cases, recurrent glioblastoma, non-surgery or surgery of enhancing tumor portion alone, biopsy-only or articles without survival data were excluded. The study was performed and presented according to the PRISMA guidelines.[7]

Data collection

Main outcome of interest was overall survival as hazard ratio (HR) with 95% confidence interval (CI) and median overall survival with 95% CI of both GTR and SPTR groups. These data were collected from the included studies or calculated based on other available data or extracted based on data points created by a pixel-by-pixel method from survival curves. GTR and SPTR were defined by the authors of the studies that were included in our meta-analysis. We categorized GTR as 100% resection of CE and SPTR as every effort of resection beyond GTR of CE, which was qualitatively or quantitatively defined by authors of the included studies. Li et al. presented their data on survival and neurological outcome mainly by categorizing <53 % FLAIR resection and ≥53% FLAIR resection. However, after email correspondence with the senior author, we received additional survival data on the GTR and SPTR groups to perform our meta-analysis.[8]

When available, data on post-operative new neurological deficits, surgical complications, or quality of life after surgery was collected to assess the safety of both SPTR and GTR.

Meta-analysis

HR with 95% CI and median overall survival with 95% CI of both SPTR and GTR groups were collected or calculated for each study based on available data using the random effects model and presented in forest plots using Review Manager (RevMan, version 5.3; Cochrane Collaboration). The random effects model was used instead of the fixed effects model due to heterogeneity between the studies, in order to give a more conservative and a clinically

reliable interpretation of the summarized statistics and the confidence intervals. The HRs were corrected by the authors of the individual papers for several prognostic factors across studies such as age, Karnofsky Performance Status (KPS), adjuvant therapy, tumor volume and location and methylguanine-DNA methyltransferase (MGMT) methylation and isocitrate dehydrogenase (IDH) mutation status (see Supplementary Materials for details). A p value of < 0.05 was considered to be statistically significant. Heterogeneity was calculated and interpreted using the Chi-squared test and I^2 values with Review Manager.

Quality of evidence

Quality of evidence obtained from the articles in this study was graded based on the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) system.[9] Within this system, four levels of quality rating can be assigned to studies; from 'high', often given to randomized controlled trials, to 'moderate', 'low' often given to observational studies, and 'very low'. The quality of evidence was rated by F.I. and was based on methodological quality, risk of biases (based on ROBINS-I),[10] heterogeneity, and precision of effect estimates.

Results

We identified 1,796 citations from Embase, 1,503 from Pubmed, 1,424 from Web of Science, 140 from Cochrane Central and 200 from Google Scholar. Removal of duplicate articles resulted in 2,346 unique citations that were screened based on title and abstracts. In total, 2,322 off-topic articles were excluded, leaving 24 articles that were read in full. [6,8,11-32] After reading, 18 articles were excluded, due to absence of SPTR (N=11), absence of sufficient survival data (N=4), and article type (N=3; two commentaries and one review). This resulted in a final total of six studies [8,12,18,19,21,28] that were included in our systematic review and meta-analysis. The process of study inclusion is presented in the PRISMA flow diagram in Figure 1 and demographics of the included studies are presented in Table 1.

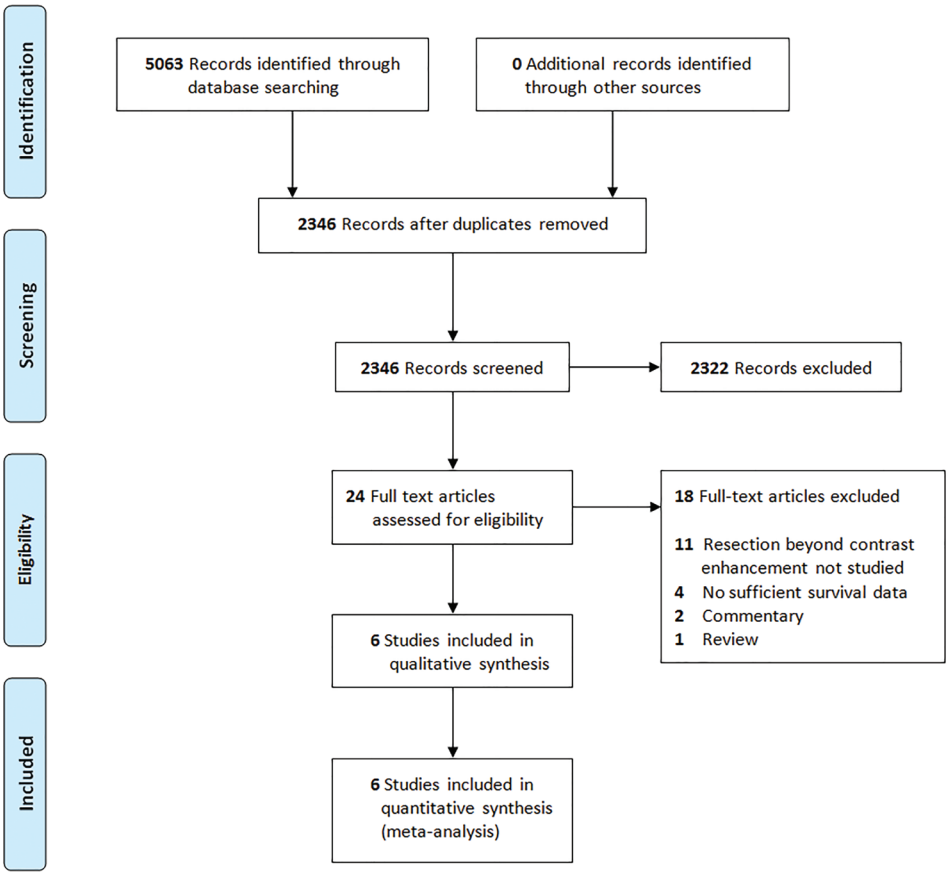


Figure 1. PRISMA Flow Diagram.

Table 1. Study demographics.

Study	Country	GTR defined as resection of	SPTR defined as resection of	HR: group comparison*	HR: multivariate analysis†	Quality of evidence (GRADE)‡
Aldave 2013[12]	Spain	100% of CE	100% of CE + total resection of fluorescing tumor	Yes	Yes	3
Li 2016[8]	United States	100% of CE	100% of CE + >0% to 100% FLAIR	Yes	Yes	2
Eyupoglu 2016[19]	Germany	100% of CE	100% of CE + total resection of fluorescing tumor	Yes	No	3
Pessina 2017[28]	Italy	100% of CE	100% of CE + 100% of FLAIR	Yes	Yes	3
Esquanazi 2017[18]	United States	95-100% of CE	>100% of CE	Yes	Yes	3
Glenn 2018[21]	United States	100% of CE	100% of CE + > 1 cm surrounding brain tissue	Yes	Yes	4

GTR: Gross total resection, SPTR: Supra total resection, CE: Contrast enhancement, HR: Hazard Ratio * Yes: HR should be interpreted as a between group comparison, No: HR should be interpreted as per unit of residual T2-FLAIR volume

† Which included prognostic factors in the model such as age, tumor volume, tumor location, Karnofsky Performance Status, MGMT methylation status and adjuvant therapy. For a full overview, see Supplementary Materials.

‡ Levels of quality: 1: high, 2: moderate, 3: low, 4: very low. For details, see Supplementary Materials.

Meta-analysis

Our meta-analysis based on six studies comparing SPTR with GTR [8,12,18,19,21,28] included 1,168 unique patients with GBM and showed that when compared with GTR, SPTR of GBM resulted in 53% lower risk of mortality at any time during follow up (HR = 0.47, 95% CI = 0.31 - 0.72, P = 0.0005, I^2 = 68%) (Figure 2).

The median overall survival for GTR and SPTR was 15.0 months (95% CI = 2.9 – 26.1) and 28.3 months (95% CI = 12.5- 44.1), respectively (P < 0.0001). The survival benefit with SPTR was 13.3 months (95% CI = 3.6 - 23.1) when compared with GTR (P = 0.007). However, Esquanazi et al.[18] especially influenced the heterogeneity (I^2) of the median overall survival analysis, possibly due to a much higher median overall survival for SPTR of 54 months within this study. Exclusion of this study in an additional median overall survival analysis showed a decrease of I^2 from 96% to 50%. When compared with GTR, SPTR of GBM resulted in a longer median overall survival of 6.4 months (95% CI = 3.2 - 9.7, P = 0.0001) (Figure 3).

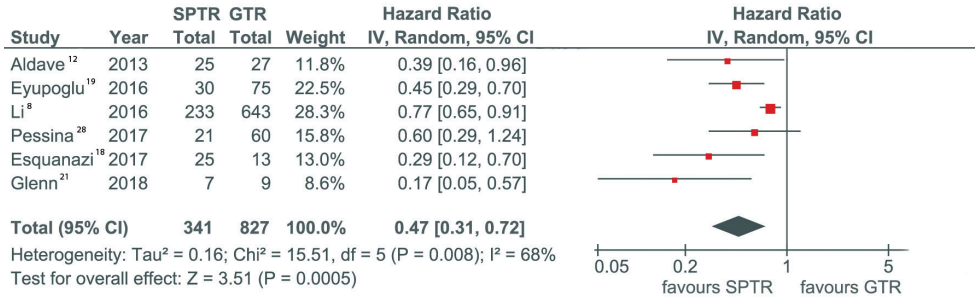


Figure 2. Forest plot with hazard ratio and 95% confidence interval (CI) for supratotal resection (SPTR) versus gross total resection (GTR).

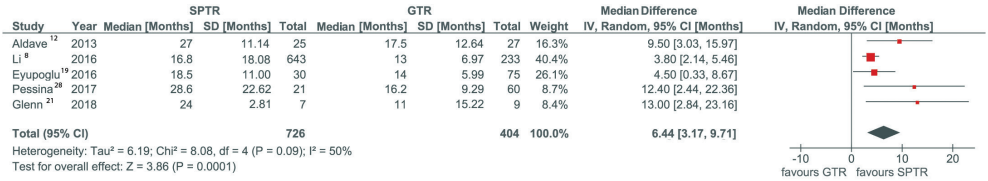


Figure 3. Forest plot with median overall survival difference between supratotal resection (SPTR) versus gross total resection (GTR).

Quality of evidence

No randomized controlled trials were available for inclusion in this meta-analysis; all studies were retrospective with a non-randomized treatment assignment. Following the GRADE guidelines, studies were assigned as 'low' level [12,18,19,28] quality evidence, upgraded to a 'moderate' level[8] or downgraded to a 'very low' level [21] (Table 1) based on previously mentioned mainly methodological factors such as study design, internal validity, risk of biases, and precision (for details; see Supplementary Materials).

Post-operative new neurological deficits

While in three studies incidence rates of post-operative new neurological deficits for both SPTR (0%-19%) and GTR (0%-8%) groups were reported, in other three studies this data was either not (clearly) presented or absent (Table 2). Overall, reports were limited and based on this data the difference in post-operative neurological outcome between SPTR vs GTR could not clearly be analyzed.

Table 2. Reported incidence rates of post-operative neurological deficits.

Study	GTR N (%)	SPTR N (%)	P value	Comments
Aldave 2013 [12]	2 (8)	4 (18.5)	0.27	No details on the deficits was reported.
Li 2016 [8]	NA	120 (19)	NA	Motor deficits, speech and visual impairments were the most commonly reported 30-day post-operative neurological complications.
Glenn 2018 [21]	0 (0)	1 (11.1)	0,66	One patient with dysphasia.
Pessina 2017 [28]	0 (0)	0 (0)	NA	There were no patients that developed post-operative new neurological deficits in the GTR and SPTR groups.
Eyupoglu 2016 [19]	see comments		0.47 – 1.0	Rates on post-operative new neurological deficits between groups was not clearly collectable. Overall no significant worsening of motor, visual, speech, cognitive deficits or seizures was reported in both groups.
Esquanazi 2017 [18]	NA	NA	NA	3 (7.9%) had transient and 2 (5.3%) had permanent neurological deficits of all 38 patients. Group-specific data was not reported.

NA: not applicable.

Discussion

This PRISMA guided systematic review and meta-analysis studied the association between SPTR and survival in patients with GBM. This study based on six articles and 1,168 unique patients showed that when compared with GTR, SPTR of GBM resulted in 53% lower risk of mortality at any time during follow-up, and 6.4 months longer median overall survival. The quality of evidence of the available studies, however, was moderate to very low. Furthermore, the incidence rates of post-operative new neurological deficits were minimally and only qualitatively reported, and data on quality of life were not reported in any of the included studies. It therefore remains unclear whether SPTR can be achieved as safely as GTR.

Defining SPTR

There are a multitude of definitions for SPTR. The term was first coined by Yordanova et al. in 2011 to explain the procedure of low grade glioma resection guided by functional boundaries, defined by the area of T2-FLAIR hyperintensity [33] during awake surgery, instead of classical neuronavigation based pre-operative MRI post-contrast scans. The authors showed that SPTR of low-grade glioma resulted in significant delay of anaplastic degeneration (P = 0.037) without causing permanent neurological deficits.

With SPTR of GBM, the classically used CE boundaries of the tumor are exceeded to include the non-enhancing hyperintense T2-FLAIR region in the resection, presumed to represent edematous, tumor infiltrated tissue. This region can infiltrate deep into eloquent brain areas[3]. For our analysis, we considered any effort of resection beyond GTR of CE as SPTR, irrespective of the procedure that was used to achieve this. While some studies had a relatively conservative SPTR procedure such as resection of additional residual 5-ALA fluorescent tissue [12,19] or resection of an additional rim (>1cm) after GTR of CE [21], Li et al. and Pessina et al. defined SPTR respectively as additional FLAIR abnormality resection (range >0% to 100%)[8] and as 100% resection of T2-FLAIR hyperintense volumes in addition to GTR of CE [28]. Additionally, a very recent study defined SPTR differently as GTR plus (frontal or temporal) lobectomy, without any volumetric analysis on T2-FLAIR volumes or functionally defined borders during awake surgery[34]. Therefore, based on literature, SPTR can only broadly be defined as every effort of resection after GTR of CE portion of GBM; no standardized definition of SPTR exists and thus needs to be defined in a prospective setting.

Li et al. highlighted that in comparison with GTR, the relative low rates of post-operative new neurological deficits in the SPTR group could probably be explained by the increased use of awake surgery with neurophysiological monitoring, (sub)cortical mapping and imaging guidance such as intra-operative MRI, fMRI and DTI navigation.[2,8] Although Jenkinson et al. showed in a recent Cochrane review that intra-operative MRI or 5-ALA fluorescent guided surgery may help to increase the extent of GBM resection,

reports on adverse events were incomplete and studies had very low quality evidence.[2] Since portions of hyperintense T2-FLAIR area may represent both tumor infiltrated edematous and solely edematous portions, additional methods to delineate non-CE tumor infiltration are important to obtain a target for SPTR. Resection of the former may contribute to a better patient survival, while the impact of the latter on survival will be limited, but would pose an unnecessary risk of complications. In addition to 5-ALA, physiological pre-operative MRI imaging such as positron emission tomography (PET) MRI may help to identify tumor infiltrated, high metabolic portions outside the CE areas of GBM as a specific target during surgery.[35]

Limitations

Overall, the results of our analysis should be interpreted with caution due to several limitations. First, because of between study heterogeneity (I^2=68%) across analyzed studies, which possible arose due to several factors; the differences in defining SPTR; the heterogeneous or limited sizes of small study populations across studies; and their retrospective design and non-randomized treatment assignment. Indeed, as mentioned earlier, different definitions of SPTR across studies, which are often retrospectively defined and which currently varies between additional resection of fluorescing tumor to total resection of T2-FLAIR hyperintensity on top of GTR of CE, makes it challenging to draw a clear conclusion that would benefit clinical practice of neurosurgeons. Secondly, since the survival benefit and safety of resecting T2-FLAIR hyperintense regions is unclear, there was no clinical equipoise to randomize GBM patients into SPTR or GTR groups. Therefore, SPTR was probably performed in only a selected group of patients, since pushing the limits of GBM resection beyond CE was often not the main goal of surgery, especially not in tumors located in eloquent brain area.

Finally, a heterogeneous patient population with different prognostic factors such as age, KPS, adjuvant therapy, IDH mutation and MGMT methylation status, tumor volume and location were present across studies. Although corrections for some of these factors were applied in multivariate analyses across studies, the real impact of these factors on achieving SPTR and on survival are difficult to determine due to the retrospective nature of the studies. GBM location and eloquence is a known important prognostic factor. A very recent study suggests that GBM located in non-eloquent area such as the frontal or temporal pole could possibly receive safe SPTR with a total lobectomy.[34] However, no clear comparison with GBM located in eloquent area was available. In the articles we have included in our study, correction for GBM eloquence was performed only limited (as presented in Supplementary Materials). Furthermore, some of these factors such as genetic mutation, location and tumor size are inter correlated and their influence on survival is difficult to investigate retrospectively; GBM with IDH mutation (and MGMT methylation) have a significant better survival and at the same time are more often located in the frontal pole and have larger tumor size.[36]

These limitations introduced biases such as confounding by indication, information and selection biases in the analyzed studies, which consequently have influenced internal validity and the quality of the studies negatively. Indeed, since the quality of evidence of the analyzed studies were only moderate to very low, the benefit of SPTR of GBM on survival and its safety remains unclear and at this point of time and cannot be considered as a main goal of GBM surgery. Taking these limitations into account, we used a random effect model to provide a conservative, but clinically more reliable and meaningful interpretation of the treatment effect of SPTR [37]. The random effects model summary results of HR 0.47 (95% CI = 0.31 - 0.72) and median overall survival benefit of 6.4 months (95% CI % 3.2 - 9.7) provides a more reliable average estimate effect because it takes real differences in the SPTR versus GTR effect in each study into account and gives wider uncertainty around the estimate, as compared with fixed effects model summary results of HR 0.67 (95% CI = 0.58 - 0.78) and a median overall survival benefit of 4.6 months (95% CI = 3.9 – 6.0).

It should be underlined that SPTR should not be attempted at any cost. As always, the aim of GBM surgery remains to maximize the extent of GBM resection without causing new neurological deficits and by maintaining a good quality of patient's life. Any potential benefit of a prolonged survival by extending tumor resection could be considered by exploring (functional) boundaries during awake surgery and additional imaging adjuncts such intra-operative MRI, ultrasound and 5-ALA. This potential benefit should be weighed, together with the patient, against the risk of pushing the boundaries of GBM resection and its effect on post-operative neurological functioning and quality of life.

In our study, the incidence on post-operative new neurological deficits was not well reported across different studies. Post-operative neurological status was reported only descriptively, without the use of validated tests and quantitative outcome measures. Additionally, no data on quality of patient lives were reported across the studies. The safety of the procedure can therefore not be established.

Future Directions

Future well designed, prospective, randomized controlled trials are thus needed to investigate both the safety and survival benefit of SPTR of GBM. These trials should include prospective, and clearly defined volumetric SPTR measures in order to standardize the definition of SPTR. Such a standardized definition of SPTR is expected to improve comparability between studies and improve the quality of evidence of studies investigating SPTR of GBM. Advanced physiological imaging techniques could additionally be used to identify non-CE infiltrating, residual tumor beyond the CE margins. Finally, validated measures on post-operative neurological outcome and quality of life assessment (QLQ-C30/QLQ-BN20) corrected for important prognostic factors such as age, KPS, adjuvant therapy, IDH mutation, and MGMT methylation status should be used to study the benefit of SPTR of GBM on patient survival and safety.[38]

Conclusion

To the best of our knowledge, we performed the first meta-analysis on SPTR of GBM, suggesting that when compared with GTR, SPTR of GBM shows lower risk of mortality at any time during follow-up and longer overall median survival. These results should however be interpreted with caution, as the quality of the available evidence was only moderate to very low, and importantly, there was no consensus on the definition of SPTR and no reliable information on the safety of SPTR was available. Therefore, based on current available literature, the benefit of SPTR of GBM on survival and its safety remains unclear at this point of time and can as yet not be considered as standard of care for GBM surgery. Our findings however, does encourage further investigation in well-designed, prospective, randomized trials to clarify whether SPTR can be achieved safely to improve survival of patients with GBM.

References

1. Brown TJ, Brennan MC, Li M, et al. Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis. *JAMA Oncol.* 2016;2(11):1460-1469.
2. Jenkinson MD, Barone DG, Bryant A, et al. Intraoperative imaging technology to maximise extent of resection for glioma Review. *Cochrane Database Syst Rev.* 2018;1:CD012788.
3. Eidel O, Burth S, Neumann JO, et al. Tumor Infiltration in Enhancing and Non-Enhancing Parts of Glioblastoma: A Correlation with Histopathology. *PLoS One.* 2017;12(1):e0169292.
4. Díez Valle R, Tejada Solis S, Idoate Gastearena MA, García De Eulate R, Domínguez Echávarri P, Aristu Mendiroz J. Surgery guided by 5-aminolevulinic fluorescence in glioblastoma: Volumetric analysis of extent of resection in single-center experience. *J Neuro-Oncol.* 2011;102(1):105-113.
5. Duffau H. A new philosophy in surgery for diffuse low-grade glioma (DLGG): Oncological and functional outcomes. *Neurochirurgie.* 2013;59(1):2-8.
6. Duffau H. Is supratotal resection of glioblastoma in noneloquent areas possible? *World Neurosurg.* 2014;82(1-2):E101-E103.
7. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Bmj.* 2009;339:b2535.
8. Li YM, Suki D, Hess K, Sawaya R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? *J Neurosurg.* 2016;124(4):977-988.
9. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj.* 2008;336(7650):924-926.
10. Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *Bmj.* 2016;355:i4919.
11. Abd-El-Barr MM, Chiocca EA. How Much Is Enough? The Question of Extent of Resection in Glioblastoma Multiforme. *World Neurosurg.* 2014;82(1-2):e109-e110.
12. Aldave G, Tejada S, Pay E, et al. Prognostic value of residual fluorescent tissue in glioblastoma patients after gross total resection in 5-aminolevulinic acid-guided surgery. *Neurosurgery.* 2013;72(6):915-920.
13. Bahrami N, Piccioni D, Karunamuni R, et al. Edge contrast of the FLAIR hyperintense region predicts survival in patients with high-grade gliomas following treatment with bevacizumab. *Am J Neuroradiol.* 2018;39(6):1017-1024.
14. Catapano G, Sgulò FG, Seneca V, Lepore G, Columbano L, di Nuzzo G. Fluorescein-Guided Surgery for High-Grade Glioma Resection: An Intraoperative "Contrast-Enhancer". *World Neurosurg.* 2017;104:2-5.
15. Coburger J, Engelke J, Scheuerle A, et al. Tumor detection with 5-aminolevulinic acid fluorescence and Gd-DTPA-enhanced intraoperative MRI at the border of contrast-enhancing lesions: a prospective study based on histopathological assessment. *Neurosurg Focus.* 2014;36(2):E3.

16. Coburger J, Hagel V, Wirtz CR, König R. Surgery for glioblastoma: Impact of the combined use of 5-aminolevulinic acid and intraoperative MRI on extent of resection and survival. *PLoS ONE.* 2015;10(6).
17. Elson A, Bovi J, Siker M, Schultz C, Paulson E. Evaluation of absolute and normalized apparent diffusion coefficient (ADC) values within the post-operative T2/FLAIR volume as adverse prognostic indicators in glioblastoma. *J Neuro-Oncol.* 2015.
18. Esquenazi Y, Friedman E, Liu Z, Zhu JJ, Hsu S, Tandon N. The Survival Advantage of "supratotal" Resection of Glioblastoma Using Selective Cortical Mapping and the Subpial Technique. *Neurosurgery.* 2017;81(2):275-288.
19. Eyüpoglu IY, Hore N, Merkel A, Buslei R, Buchfelder M, Savaskan N. Supra-complete surgery via dual intraoperative visualization approach (DiVA) prolongs patient survival in glioblastoma. *Oncotarget.* 2016;7(18):25755-25768.
20. Gessler F, Forster MT, Duetzmann S, et al. Combination of Intraoperative Magnetic Resonance Imaging and Intraoperative Fluorescence to Enhance the Resection of Contrast Enhancing Gliomas. *Neurosurgery.* 2015;77(1):16-22.
21. Glenn CA, Baker CM, Conner AK, et al. An Examination of the Role of Supramaximal Resection of Temporal Lobe Glioblastoma Multiforme. *World Neurosurg.* 2018;114:e747-e755.
22. Grabowski MM, Recinos PF, Nowacki AS, et al. Residual tumor volume versus extent of resection: Predictors of survival after surgery for glioblastoma. *J Neurosurg.* 2014;121(5):1115-1123.
23. Grossman R, Shimony N, Shir D, et al. Dynamics of FLAIR Volume Changes in Glioblastoma and Prediction of Survival. *Ann Surg Oncol.* 2017;24(3):794-800.
24. Mampre D, Ehresman J, Pinilla-Monsalve G, et al. Extending the resection beyond the contrast-enhancement for glioblastoma: feasibility, efficacy, and outcomes. *Br J Neurosurg.* 2018:1-8.
25. Marko NF, Weil RJ, Schroeder JL, Lang FF, Suki D, Sawaya RE. Extent of resection of glioblastoma revisited: Personalized survival modeling facilitates more accurate survival prediction and supports a maximum-safe-resection approach to surgery. *J Clin Oncol.* 2014;32(8):774-782.
26. Meyer RB, Bates LM, Goerss SJ, et al. Awake craniotomy for aggressive resection of primary gliomas located in eloquent brain. *Mayo Clinic Proceedings.* 2001;76(7):677-687.
27. Neira JA, Ung TH, Sims JS, et al. Aggressive resection at the infiltrative margins of glioblastoma facilitated by intraoperative fluorescein guidance. *J Neurosurg.* 2017;127(1):111-122.
28. Pessina F, Navarria P, Cozzi L, et al. Maximize surgical resection beyond contrast-enhancing boundaries in newly diagnosed glioblastoma multiforme: is it useful and safe? A single institution retrospective experience. *J Neuro-Oncol.* 2017;135(1):129-139.
29. Quan GM, Zheng YL, Yuan T, Lei JM. Increasing FLAIR signal intensity in the postoperative cavity predicts progression in gross-total resected high-grade gliomas. *J Neuro-Oncol.* 2018;137(3):631-638.
30. Schucht P, Knittel S, Slotboom J, et al. 5-ALA complete resections go beyond MR contrast enhancement: Shift corrected volumetric analysis of the extent of resection in surgery for glioblastoma. *Acta Neurochir.* 2014;156(2):305-312.

31. Yan JL, Van Der Hoorn A, Larkin TJ, Boonzaier NR, Matys T, Price SJ. Extent of resection of peritumoral diffusion tensor imaging-detected abnormality as a predictor of survival in adult glioblastoma patients. *J Neurosurg.* 2017;126(1):234-241.
32. Yordanova YN, Duffau H. Supratotal resection of diffuse gliomas – an overview of its multifaceted implications. *Neurochirurgie.* 2017;63(3):243-249.
33. Yordanova YN, Moritz-Gasser S, Duffau H. Awake surgery for WHO Grade II gliomas within "noneloquent" areas in the left dominant hemisphere: toward a "supratotal" resection. Clinical article. *J Neurosurg.* 2011;115(2):232-239.
34. Roh TH, Kang SG, Moon JH, et al. Survival benefit of lobectomy over gross-total resection without lobectomy in cases of glioblastoma in the noneloquent area: a retrospective study. *J Neurosurg.* 2019:1-7.
35. John F, Bosnyak E, Robinette NL, et al. Multimodal Imaging-defined Subregions in Newly-diagnosed Glioblastoma: Impact on Overall Survival. *Neuro Oncol.* 2018.
36. Smits M, van den Bent MJ. Imaging Correlates of Adult Glioma Genotypes. *Radiology.* 2017;284(2):316-331.
37. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *Bmj.* 2011;342:d549.
38. Gessler F, Bernstock JD, Braczynski A, et al. Surgery for Glioblastoma in Light of Molecular Markers: Impact of Resection and MGMT Promoter Methylation in Newly Diagnosed IDH-1 Wild-Type Glioblastomas. *Neurosurgery.* 2018.

Supplementary materials

Search queries for several literature sources

Embase.com

('glioblastoma'/exp OR (glioblastom OR (maligna* NEAR/3 glioma*) OR (high* NEAR/3 grade* NEAR/3 glioma*) OR ((grade-iv OR grade-4) NEAR/3 glioma*) OR gbm):ab,ti,kw) AND ('surgery'/de OR 'neurosurgery'/exp OR 'cancer surgery'/exp OR 'brain surgery'/exp OR 'brain tumor'/exp/dm_su OR 'surgical technique'/de OR (surg* OR neurosurg* OR resect*):ab,ti,kw) AND ('extent of resection'/de OR 'surgical margin'/exp OR 'fluid attenuated inversion recovery'/de OR 'fluid attenuated inversion recovery imaging'/de OR 'gross total resection'/de OR (flair OR (Fluid NEAR/3 attenuat* NEAR/3 invers* NEAR/3 recover*) OR t2 OR t-2 OR gross-total OR ((exten* OR Supratotal* OR Supramaxim*) NEAR/3 (resect* OR remov*)) OR ((surg* OR excis* OR resect*) NEAR/3 margin*) OR ((beyond OR additional*) NEAR/6 (contrast OR boundar*))):ab,ti,kw) NOT ([Conference Abstract]/lim OR [Letter]/lim OR [Note]/lim OR [Editorial]/lim)*

Web of Science

((glioblastom OR (maligna* NEAR/3 glioma*) OR (high* NEAR/3 grade* NEAR/3 glioma*) OR ((grade-iv OR grade-4) NEAR/3 glioma*) OR gbm):ab,ti,kw) AND ((surg* OR neurosurg* OR resect*):ab,ti,kw) AND ((flair OR (Fluid NEAR/3 attenuat* NEAR/3 invers* NEAR/3 recover*) OR t2 OR t-2 OR gross-total OR ((exten* OR Supratotal* OR Supramaxim*) NEAR/3 (resect* OR remov*)) OR ((surg* OR excis* OR resect*) NEAR/3 margin*) OR ((beyond OR additional*) NEAR/6 (contrast OR boundar*))):ab,ti,kw)*

Cochrane CENTRAL

((glioblastom OR (maligna* NEAR/3 glioma*) OR (high* NEAR/3 grade* NEAR/3 glioma*) OR ((grade-iv OR grade-4) NEAR/3 glioma*) OR gbm):ab,ti,kw) AND ((surg* OR neurosurg* OR resect*):ab,ti,kw) AND ((flair OR (Fluid NEAR/3 attenuat* NEAR/3 invers* NEAR/3 recover*) OR t2 OR "t 2" OR "gross total" OR ((exten* OR Supratotal* OR Supramaxim*) NEAR/3 (resect* OR remov*)) OR ((surg* OR excis* OR resect*) NEAR/3 margin*) OR ((beyond OR additional*) NEAR/6 (contrast OR boundar*))):ab,ti,kw)*

Google Scholar

glioblastoma|"malignant glioma"|"high grade glioma" flair|"Fluid attenuating inverse"|"gross total"|"Supratotal|Supramaximal resection"|"extent of resection"

GRADE quality ranking

Study	Quality level*	Underlying methodology	Risk of bias **	Precision
Aldave 2013 [12]	3	Non-random treatment allocation, retrospective	Moderate	Low N, wide CI
Li 2016 [8]	2	Non-random treatment allocation retrospective	Moderate	High N, narrow CI
Eyupoglu 2016 [19]	3	Non-random treatment allocation retrospective	Moderate	Moderate N, CI
Pessina 2017 [28]	3	Non-random treatment allocation retrospective	Moderate	Moderate N, CI
Esquanazi 2017 [18]	3	Non-random treatment allocation retrospective	Moderate	Low N, wide CI
Glenn 2018 [21]	4	Non-random treatment allocation, retrospective	Serious	Low N, wide CI

N = number of patients CI = Confidence Interval

* Levels of quality (GRADE): 1: high, 2: moderate, 3: low, 4: very low
** "Risk Of Bias In Non-randomised Studies - of Interventions" (ROBINS-I), based on risks of selection, information and confounding biases, with summarized judgements which includes moderate, serious and critical risks of bias or no information.

Overview of multivariate analyses across studies with adjustment for prognostics factors.

10.1

Study	Multivariate analysis	Prognostic factors as variables
Aldave 2013 [12]	Yes	age, KPS, MGMT methylation status, tumor eloquent location, preoperative tumor volume, and adjuvant therapy.
Li 2016 [8]	Yes	unclear, at least treatment history
Eyupoglu 2016 [19]	No	Not performed
Pessina 2017 [28]	Yes	age, extent of contrast enhancement resection, tumor location
Esquanazi 2017 [18]	Yes	age, tumor volume, KPS, intra operative mapping, BCNU wafer
Glenn 2018 [21]	Yes	MGMT methylation status, tumor laterality

KPS = Karnofsky Performance Status, MGMT= methylguanine-DNA methyltransferase IDH= isocitrate dehydrogenase, BCNU= bis-chloroethylnitrosourea or Carmustine.

Chapter 10.2

Letter to the Editor: Supratotal resection of glioblastoma

Fatih Incekara, Marion Smits, Arnaud J.P.E. Vincent

J Neurosurg. 2019 Aug 16:1-2

Supratotal resection of glioblastoma

TO THE EDITOR: We read with great interest the research by Roh et al. (Roh TH, Kang SG, Moon JH, et al: Survival benefit of lobectomy over gross-total resection without lobectomy in cases of glioblastoma in the noneloquent area: a retrospective study. J Neurosurg 2019 Mar 1:1-7. doi: 10.3171/2018.12.JNS182558. [Epub ahead of print]) regarding supratotal glioblastoma (GBM) resection in non-eloquent brain area.[9]

In this well written and timely article, the authors show that in patients with GBM located in right sided, frontal and temporal, non-eloquent brain area, supratotal resection (SupTR) defined as a total lobectomy improves median OS and PFS, when compared with solely gross total resection (GTR) of GBM. We commend the authors for critically appraising their clinical practice variations, which adds to our understanding on the potential value of SupTR of GBM in improving patient survival.

The study however, does raise some issues which need careful consideration. First, studying the value of SupTR without detailed and preferably standardized assessment of neurological outcome (i.e. NIHSS)[10] and quality of life assessments (i.e. QLQ-C30 and QLQ-BN20)[11] remains an important limitation. Although this issue is acknowledged by the authors, it is only briefly discussed in the limitation paragraph. Readers should be more strongly notified to not draw any decisive conclusions on the safety of this procedure even in noneloquent area, since only KPS was measured as a global neurological outcome and no data was available on the quality of patient life. Within this context, it also remains unclear how the dominant hemisphere and eloquent brain area were determined. The exclusion of all-left sided tumors by definition suggests that no functional MRI was used to asses functional lateralization and seems to disregard the notion that the right hemisphere may be dominant in some patients. Additionally, it should be noted that despite the exclusion of "tumors involving motor or language areas", awake surgery was performed when GBM "was near an eloquent area". We think that it would be beneficial to read in Table 1 how many cases were classified as (near) eloquent GBM and thus how many cases were performed with awake surgery. If this, hypothetically, was the case in the majority of patients, many noneloquent GBM should in fact have been labeled as (near) eloquent located GBM. In such a case, the conclusion that SupTR of noneloquent GBM is a safe procedure would - unintentionally - mislead the reader.

Secondly, as the authors acknowledge, the definition of supratotal resection is ambiguous. GBM is known to infiltrate beyond the margins of contrast enhancement, into the surrounding edematous T2/FLAIR-weighted hyperintense area, as seen on MRI.[2] In this study, SupTR was defined as complete resection of contrast enhancement as seen on T1-weighted postcontrast, postoperative MRI plus (frontal or temporal) lobectomy. In

addition, the authors considered a resection cavity on post-operative MRI that was larger than the contrast enhancing tumor volume on pre-operative MRI as a criterion for SupTR. We think that this criterion is highly debatable because of two reasons. First, factors such as brain shift and swelling of brain parenchyma makes a comparison of post-surgical cavity with pre-operative tumor volume unreliable. Secondly, as we mentioned earlier, GBM can infiltrate far beyond contrast enhancement and in fact, beyond the resection cavity as seen on post-contrast T1-weighted images, due to expansive and large T2/FLAIR-weighted hyperintense regions. The authors do not specify whether SupTR included the entire T2/FLAIR hyperintense region. In fact, unfortunately, no volumetric analysis on T2/FLAIR abnormality was presented in this study. Previously published studies investigating the role of supratotal resection or post-operative residual T2/FLAIR abnormalities found an association with survival.[1,3-8] We believe that currently no clear definition of supratotal resection exists and it can only broadly be defined as any effort of resection beyond the contrast enhancing part of GBM. In future studies it yet needs to be defined in a prospective setting.

Finally, we fear that the internal and external validity of the study is influenced by respectively confounding and selection biases. Although the authors acknowledge the presence of selection bias in the last sentence of the discussion, this issue needs more attention, since the impact of the biases on the results and the main conclusion is expected to be substantial. Indeed, out of 400 GBM patients only 40 patients with GTR of right-sided frontal or temporal noneloquent GBM were selected. It is also unclear how GBM location *within* a lobe and surgeon preference influenced SupTR vs GTR. The surgeon might have intended a SupTR for a GBM in the right temporal pole, whereas a more posterior and medial GBM within the same lobe might receive GTR instead of a lobectomy. Based on this highly selected group of patients, it remains unclear how to translate the potential benefit and the safety of SupTR for GBM surgery in general. In addition, the retrospective nature of the study prohibits a reliable comparison of median OS and PFS between the SupTR and GTR groups and raises several questions such as the impact of surgeon's preference on achieving SupTR vs GTR, its influence on survival and the impact of performing 5-ALA fluorescence guided or awake surgery on survival. It is therefore preferable to control for known confounding factors in a multivariate analysis and focus on the corrected hazard ratio of SupTR vs GTR, as the authors rightly performed, instead of a biased improvement of median OS and PFS between groups as seen in the Kaplan-Meier curves.

We hope that emphasizing these issues will encourage readers to carefully consider both the potential benefits *and* the risks of SupTR, such that the conclusions of this important study can be fully and reliably appreciated.

References

1. Aldave G, Tejada S, Pay E, Marigil M, Bejarano B, Idoate MA, et al: Prognostic value of residual fluorescent tissue in glioblastoma patients after gross total resection in 5-aminolevulinic Acid-guided surgery. **Neurosurgery 72:**915-920; discussion 920-911, 2013
2. Eidel O, Burth S, Neumann JO, Kieslich PJ, Sahm F, Jungk C, et al: Tumor Infiltration in Enhancing and Non-Enhancing Parts of Glioblastoma: A Correlation with Histopathology. **PLoS One 12:**e0169292, 2017
3. Esquenazi Y, Friedman E, Liu Z, Zhu JJ, Hsu S, Tandon N: The Survival Advantage of "Supratotal" Resection of Glioblastoma Using Selective Cortical Mapping and the Subpial Technique. **Neurosurgery 81:**275-288, 2017
4. Eyupoglu IY, Hore N, Merkel A, Buslei R, Buchfelder M, Savaskan N: Supra-complete surgery via dual intraoperative visualization approach (DiVA) prolongs patient survival in glioblastoma. **Oncotarget 7:**25755-25768, 2016
5. Glenn CA, Baker CM, Conner AK, Burks JD, Bonney PA, Briggs RG, et al: An Examination of the Role of Supramaximal Resection of Temporal Lobe Glioblastoma Multiforme. **World Neurosurg 114:**e747-e755, 2018
6. Kotrotsou A, Elakkad A, Sun J, Thomas GA, Yang D, Abrol S, et al: Multi-center study finds postoperative residual non-enhancing component of glioblastoma as a new determinant of patient outcome. **J Neurooncol 139:**125-133, 2018
7. Li YM, Suki D, Hess K, Sawaya R: The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? **J Neurosurg 124:**977-988, 2016
8. Pessina F, Navarria P, Cozzi L, Ascolese AM, Simonelli M, Santoro A, et al: Maximize surgical resection beyond contrast-enhancing boundaries in newly diagnosed glioblastoma multiforme: is it useful and safe? A single institution retrospective experience. **J Neurooncol 135:**129-139, 2017
9. Roh TH, Kang SG, Moon JH, Sung KS, Park HH, Kim SH, et al: Survival benefit of lobectomy over gross-total resection without lobectomy in cases of glioblastoma in the noneloquent area: a retrospective study. **J Neurosurg:**1-7, 2019
10. Stummer W, Pichlmeier U, Meinel T, Wiestler OD, Zanella F, Reulen HJ, et al: Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. **Lancet Oncol 7:**392-401, 2006
11. Taphoorn MJ, Stupp R, Coens C, Osoba D, Kortmann R, van den Bent MJ, et al: Health-related quality of life in patients with glioblastoma: a randomised controlled trial. **Lancet Oncol 6:**937-944, 2005

Chapter 11

Summary & Conclusion

Summary and conclusion

This thesis assessed the value of glioblastoma imaging and resection in light of molecular markers. Here, I summarize and discuss the main findings in this thesis. I provide future perspectives on topics addressed in this thesis and I finish with a conclusion.

In **Part I**, we predict molecular markers of glioma, primarily IDH mutation, 1p/19q codeletion and MGMT promoter methylation, and cognitive outcome of patients based on preoperative MRI scans. In **Part II**, we present the results of a randomized controlled trial that assesses the value of intraoperative ultrasound guided surgery on the extent of glioblastoma resection. We answer the question whether intraoperative ultrasound guided surgery enables complete tumor resection more often than standard glioblastoma surgery. Based on the results, we provide a recommendation to achieve maximal safe resection. In addition, we provide a proof of concept on the use of wearable mixed reality device for glioblastoma localization on the operating room. In **Part III** of this thesis, we study the association between the extent of contrast-enhancing and non-contrast enhancing tumor resection and survival of patients with glioblastoma. We validate our findings in a multi-center glioblastoma cohort. In addition, we assess the value of supratotal resection in glioblastoma, that is, tumor resection beyond the borders of contrast-enhancement.

Main findings

Part I: Preoperative imaging

In **Chapter 2**, we re-evaluated the topographical distribution of 436 glioblastoma with vs. without MGMT promoter methylation using voxel-wise MRI analysis in one of the largest homogenous IDH wildtype glioblastoma populations. Visual inspection indicated that when compared with MGMT unmethylated glioblastoma, methylated glioblastoma were more frequently located near bifrontal and left occipital periventricular area and less frequently near the right occipital periventricular area. These visual appearance of differences, however, could not be confirmed with rigorous voxel-wise statistical testing. We therefore conclude that there was no statistical difference in anatomical localization between IDH wildtype glioblastoma with vs. without MGMT promoter methylation. Prior to the WHO 2016 classification update,(1) studies on this topic had reported conflicting results; some studies showed that glioblastoma with MGMT methylation were lateralized to the left temporal lobe, while other showed right hemisphere lateralization.(2-5) There were also studies that reported no differences in localization, in concordance with the findings of our study.(6-9) These conflicting results in literature may partially be explained by variations in glioblastoma patient populations across studies, many of which were

performed in the pre-WHO 2016 classification era when the impact of molecular subtyping of glioblastoma according to IDH mutation status was less of a consideration.

In **Chapter 3**, we predicted the 1p/19q codeletion status in 284 low grade glioma patients on preoperative MRI scans with machine learning. We externally validated our algorithm on an independent dataset (n=129) and we compared the performance of our algorithm with that of clinical experts. We found that tumor location, heterogeneity on T2-weighted MRI scans, together with patient age and sex were the most important variables for our algorithm. We predicted the 1p/19q codeletion status of glioma with a performance that on average corresponded with experienced oncological neuroradiologists and outperformed oncological neurosurgeons. To the best of our knowledge, there are no prior studies predicting genetic mutations of low glioma on MRI using machine learning that compared their predictive performance with human performance.(10-13) Knowledge of the molecular profile of low grade glioma is clinically relevant, since it is recommended to consider them when timing postsurgical treatment with chemo- and/or radiotherapy. (14) Before clinical implementation in the future, more research is needed to improve the accuracy of AI algorithms, but we showed their potential to support decision making in clinical care of glioma.

In **Chapter 4**, we have developed and externally validated a deep learning algorithm that can accurately predict the IDH mutation status, 1p/19q co-deletion status, and grade of glioma, while simultaneously providing automatic glioma segmentation, based on pre-operative MRI scans. For patients in which tumor tissue cannot be obtained through surgical resection or diagnostic biopsy due to older age, poor neurological condition, or deep tumor localization, prediction of molecular markers on preoperative MRI scans may be clinically relevant in the future as an alternative.

In **Chapter 5**, we studied the association between the cognition of patients with a low grade glioma and white matter tract changes using diffusion tensor MRI and fiber tractography. We found that cognition in terms of language deficits in repetition of speech, imprinting and attention deficits were associated with changed microarchitecture of the arcuate fasciculus and the inferior fronto-occipital fasciculus. Our results support the hypothesis that glioma may infiltrate language white matter tracts and disrupt their microarchitecture, leading to deterioration of cognition in patients with glioma. Extensive glioma resection should be balanced carefully against persevering such infiltrated white matter tracts, preferably with awake surgery and stimulation mapping.(15)

Part II: Image-guided glioblastoma surgery

In **Chapter 6**, we present the results of the ultrasound trial which aimed to assess the value of intraoperative ultrasound guided surgery on the extent of glioblastoma resection. The main research question was: *does intraoperative ultrasound guided surgery enable complete tumor resection more often than standard glioblastoma surgery*? We showed that intraoperative ultrasound enables complete tumor resection significantly more often (35%) than standard glioblastoma surgery (8%). We found that intraoperative ultrasound guidance did not harm patients in terms of neurological outcome, functional status and quality of life. Therefore, we recommended intraoperative ultrasound guided surgery, rather than standard surgery based on neuronavigation alone, to achieve complete safe tumor resection.

Complete resection of contrast-enhancing tumor has consistently been associated with longer overall survival in patients with glioblastoma.(16) Complete tumor resection is defined both qualitatively and quantitatively across studies.(17, 18) Some studies defined complete tumor resection as no residual contrast-enhancing tumor on a post-operative MRI scan, which is a stringent definition and may result in false positive assessment of residual tumor due to non-specific contrast enhancement such as due to ischemia, small vessels or unspecific reactive tissue enhancement.(19) To take interpretation variations into account, other studies also defined complete tumor resection as contrast-enhancing residual tumor <0.175 cm^3(20, 21) or as a certain resection cut-off percentage (e.g. >98%).(22) It is known that residual tumor assessment of glioblastoma has a low interobserver agreement, introducing some degree of subjectivity when distinguishing contrast-enhancing residual tumor from non-specific contrast enhancement.(23) In our trial, we defined complete tumor resection as >99% resection of contrast-enhancing tumor volume, accepting residual contrast-enhancing volume smaller than one percent to account for non-tumor related enhancement.

It is shown in two randomized controlled trials, that 5-aminolevulinic acid and intraoperative MRI guided surgery improve the extent of glioblastoma resection.(20, 21) However, an intraoperative MRI system is expensive and prolongs surgery time with approximately one hour.(21) This might explain why intraoperative MRI is still not implemented worldwide as a standard procedure during glioblastoma surgery. An alternative less-expensive, time effective technology that is used to acquire real-time imaging during brain tumor surgery is intraoperative ultrasound.(24) Real time imaging is needed, since neuronavigation systems are typically based on pre-operative MRI scans and due to brain shift, their accuracy in representing the actual situation during surgery decreases. As Jenkinson et al showed in a Cochrane review, no randomized controlled trial is ever performed to assess the value of intraoperative ultrasound to maximize tumor resection.(25) Therefore,

we initiated first such a trial and presented the results in this thesis. We recommend intraoperative ultrasound guidance as an alternative intraoperative imaging technology, just as intraoperative MRI or 5-aminolevulinic acid guided surgery (or a combination), to maximize the extent of contrast-enhancing tumor resection during glioblastoma surgery.

In **Chapter 7**, we tested in one of the first prospective clinical pilot studies, the clinical feasibility and accuracy of a wearable mixed reality device (Microsoft Hololens) for neurosurgical planning and glioblastoma localization on the operating room. We compared a wearable augmented reality headset with a standard neuronavigation system and showed in a proof of concept using quantitative outcome measures that augmented reality has a potential for glioblastoma visualization and localization during neurosurgical planning. We observed that manual hologram placement on the patient was a limitation decreasing the accuracy of tumor localization of the augmented reality device. For clinical implementation, the accuracy of the augmented reality device should be improved in future research, preferably trough automated registration of the hologram with the patient and through integration of augmented reality with a neuronavigation system.

Part III: Extent of resection and survival

In **Chapter 8**, we assessed the association between contrast-enhancing and non-contrast enhancing tumor resection and survival in light of MGMT promoter methylation in 326 patients with newly diagnosed IDH wildtype glioblastoma. This study showed an association between maximal contrast-enhancing tumor resection, ≥30% non-contrast enhancing tumor resection, minimal residual contrast-enhancing tumor volume and longer overall survival. Kaplan Meier estimates indicated that extensive surgery was potentially more beneficial for patients with MGMT methylated glioblastoma. Complete contrast enhancing glioblastoma resection is consistently associated with patient survival.(16) This thesis however, re-evaluates this association based on two new insights.

First, previous studies reported limited data on molecular characteristics of glioblastoma (i.e. IDH mutation and MGMT promotor methylation status) since many were performed prior to the WHO 2016 classification update, in which molecular sub classification was less of a consideration.(16) Although IDH mutation within newly diagnosed primary glioblastoma is rare (<5%, and 3.5% in our cohort) these tumors represent a distinct molecular type of glioma arising from a distinct precursor lesion and the recent cIMPACT-NOW update suggests to call these tumors astrocytoma grade IV.(26-29) Therefore, new series are required to answer certain questions in tumor types defined according to this updated classification. We confirmed the association between complete contrast-enhancing tumor resection and longer survival in patients with glioblastoma IDH wildtype.

In **Chapter 9** we performed an international, multicenter, observational study including over one thousand patients with a newly diagnosed IDH-wildtype glioblastoma, in which we developed and externally validated a survival prediction model. This model included age, gender, preoperative Karnofsky performance status, extent of surgical resection, MGMT promoter methylation status, and adjuvant therapy and showed after external validation that survival of individual patients with a newly diagnosed glioblastoma could reliably be predicted. We published an online nomogram that could be useful to support shared decision making.

Secondly, this thesis re-evaluated the association between glioblastoma resection and survival by assessing the value of non-contrast enhancing tumor resection, beyond the borders of contrast enhancement. Glioblastoma is known to infiltrate far beyond the margins of contrast enhancement as seen on post-contrast T1-weighted MRI scans, into the surrounding edematous T2-weighted or T2-FLAIR hyperintense area.(30) This raises the question whether complete resection of contrast-enhancing tumor is indeed what is classically known as "gross total resection" or that resection should be extended beyond the margins of contrast enhancement, into the non-contrast enhancing area of signal abnormalities, to improve survival. The second question in this line, is which part of non-contrast enhancing tumor should be resected, as it can be hypothesized that only resection of tumor infiltrated, rather than edematous non-contrast enhanced tissue abnormalities will decrease tumor load and contribute to survival benefit. Physiological MRI is a promising tool to identify and distinguish tumor infiltrated portions from edema in non-contrast enhancing area.(31) We also showed in Chapter 8 that resection of ≥30% non-contrast enhancing tumor, beyond the margins of contrast enhancement, was associated with a longer overall survival. Other studies associated non contrast enhanced tumor resection percentages of 45%,(32) 53%(19) and 92%(33) with improved patient survival. Importantly, maximal glioblastoma resection should not be attempted at all cost and these studies reported very limited to no data on patient outcome. We also had insufficient data on patient outcome in our study, and therefore we recommended that intraoperative imaging and stimulation mapping should be used to pursue safe and maximal resection. In future research, the safety aspect of maximizing tumor resection needs to be further addressed.

The concept of supratotal resection has been known for some time within low grade glioma surgery, but it is only being investigated fairly recently in glioblastoma.(34) **Chapter 10.1** had the aim to study supratotal resection in glioblastoma surgery and its relation to patient survival in literature. We showed in our systematic review and meta-analysis based on six studies and 1168 glioblastoma patients that when compared with complete resection of contrast enhancement, supratotal resection of glioblastoma was associated with lower risk of mortality and longer median overall survival during follow

up. The quality of evidence however, was poor due to selection and confounding biases in the included studies.(19, 32, 35-38) In addition, there is much variation in the definition of supratotal resection for glioblastoma across studies. The safety aspect of supratotal resection was poorly reported, often lacking data on neurological outcome, complications or quality of patient life. Supratotal resection of glioblastoma can potentially provide survival benefit, however, future research is needed to assess its safety, as also addressed in our letter (**Chapter 10.2**).

Future perspectives

It can be expected that in future neuro-oncological radiogenomics research, the accuracy of molecular and clinical outcome predictions on MRI will increase due to 1) the rise of artificial intelligence 2) more patient data as a result of large multi-center collaborations and consortia and 3) the addition of advanced MRI techniques including physiological MRI scans such as perfusion and diffusion and combination of MRI with PET-scanning.

Physiological MRI scans will not be only of benefit for radiogenomics research, but also for safely maximizing glioma resection due to a more accurate delineation of tumor margins then that we currently are able to do based on anatomical MRI scans only. (31) This may help us to understand which part of the non-contrast enhanced portion of glioblastoma is more tumor infiltrated and thus should be resected, and which part is more edema and can be preserved.

We further expect that new intraoperative imaging technologies, combined with artificial intelligence, will be of benefit to safely maximize the extent of brain tumor resection. Already, a recent study showed potential benefit of near real-time intraoperative detection of infiltration in tumor borders within minutes using stimulated Raman spectroscopy and deep neural networks to pursue maximal and safe glioma surgery, which may replace froze sections during surgery in the future.(39) These technologies, when integrated in augmented reality navigation systems, may provide intraoperative guidance and decision-making during surgery, for example by overlaying accurate predictions of algorithms over tissue to determine whether that last part in the resection cavity is residual or healthy tumor tissue, which is often difficult to see even with a trained neurosurgical eye.

The future of neurosurgery will be shaped by new imaging technologies and artificial intelligence, and still with surgical experience, intuition and anatomical knowledge in its center, there will be a transition towards data driven surgery. It can be expected that this will not only contribute to maximize the extent of tumor resection, but the safety of brain tumor surgery will also be maintained by preserving functional brain area. Anatomical image guided surgery such as intraoperative MRI, intraoperative ultrasound and

fluorescent guided surgery will be supported by function-guided surgery such as awake surgery with stimulation mapping and intraoperative physiological image-guidance.

It can hopefully be expected, in short, that new imaging technologies and artificial intelligence will improve the way we diagnose and resect glioma, and thus eventually the way we provide clinical care for patients with brain tumors.

Conclusion

This thesis contributes to our knowledge on the value of imaging and resection of glioma, in particular glioblastoma.

We presented promising results on molecular predictions based on preoperative MRI scans using artificial intelligence. We showed in a randomized controlled trial that intraoperative ultrasound enables complete tumor resection more often than standard glioblastoma surgery, without harming patients. We recommend intraoperative ultrasound guided surgery, rather than standard surgery, to achieve complete safe tumor resection. We found that augmented reality was potentially useful to visualize and locate glioblastoma for neurosurgical planning on the operating room.

Complete glioblastoma resection was already associated with improved patient survival. We confirmed this association in molecularly defined IDHwt glioblastoma. We externally validated these results in a multi-center glioblastoma cohort and predicted survival of patients with a glioblastoma with nomograms. In addition, we found that resection of non-contrast enhanced tumor beyond contrast enhancement or supratotal resection showed potential survival benefit for patients with glioblastoma. We encourage further research on the safety of maximizing glioblastoma resection, especially beyond the borders of contrast enhancement.

Finally, we shared future perspectives, in which we expect that new imaging technologies and artificial intelligence will improve how we understand, diagnose and surgically treat patients with brain tumors.

We hope that this thesis will be of benefit for future brain tumor research, clinical care and for patients with a brain tumor.

References

1. 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 Neuropathol. 2016;131(6):803-20.
2. Ellingson BM, Lai A, Harris RJ, Selfridge JM, Yong WH, Das K, et al. Probabilistic radiographic atlas of glioblastoma phenotypes. Am J Neuroradiol. 2013;34(3):533-40.
3. Ellingson BM, Cloughesy TF, Pope WB, Zaw TM, Phillips H, Lalezari S, et al. Anatomic localization of O6-methylguanine DNA methyltransferase (MGMT) promoter methylated and unmethylated tumors: a radiographic study in 358 de novo human glioblastomas. Neuroimage. 2012;59(2):908-16.
4. Eoli M, Menghi F, Bruzzone MG, De Simone T, Valletta L, Pollo B, et al. Methylation of O6-methylguanine DNA methyltransferase and loss of heterozygosity on 19q and/or 17p are overlapping features of secondary glioblastomas with prolonged survival. Clin Cancer Res. 2007;13(9):2606-13.
5. Wang Y, Fan X, Zhang C, Zhang T, Peng X, Li S, et al. Anatomical specificity of O6-methylguanine DNA methyltransferase protein expression in glioblastomas. J Neurooncol. 2014;120(2):331-7.
6. Han Y, Yan LF, Wang XB, Sun YZ, Zhang X, Liu ZC, et al. Structural and advanced imaging in predicting MGMT promoter methylation of primary glioblastoma: a region of interest based analysis. BMC Cancer. 2018;18(1):215.
7. Carrillo JA, Lai A, Nghiemphu PL, Kim HJ, Phillips HS, Kharbanda S, et al. Relationship between tumor enhancement, edema, IDH1 mutational status, MGMT promoter methylation, and survival in glioblastoma. AJNR Am J Neuroradiol. 2012;33(7):1349-55.
8. Drabycz S, Roldán G, de Robles P, Adler D, McIntyre JB, Magliocco AM, et al. An analysis of image texture, tumor location, and MGMT promoter methylation in glioblastoma using magnetic resonance imaging. Neuroimage. 2010;49(2):1398-405.
9. Roux A, Roca P, Edjlali M, Sato K, Zanello M, Dezamis E, et al. MRI Atlas of IDH Wild-Type Supratentorial Glioblastoma: Probabilistic Maps of Phenotype, Management, and Outcomes. Radiology. 2019;293(3):633-43.
10. Chang P, Grinband J, Weinberg BD, Bardis M, Khy M, Cadena G, et al. Deep-Learning Convolutional Neural Networks Accurately Classify Genetic Mutations in Gliomas. AJNR Am J Neuroradiol. 2018;39(7):1201-7.
11. Han Y, Xie Z, Zang Y, Zhang S, Gu D, Zhou M, et al. Non-invasive genotype prediction of chromosome 1p/19q co-deletion by development and validation of an MRI-based radiomics signature in lower-grade gliomas. J Neurooncol. 2018;140(2):297-306.
12. Zhou H, Chang K, Bai HX, Xiao B, Su C, Bi WL, et al. Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas. J Neurooncol. 2019;142(2):299-307.
13. Park YW, Han K, Ahn SS, Bae S, Choi YS, Chang JH, et al. Prediction of IDH1-Mutation and 1p/19q-Codeletion Status Using Preoperative MR Imaging Phenotypes in Lower Grade Gliomas. AJNR Am J Neuroradiol. 2018;39(1):37-42.

14. van den Bent MJ, Smits M, Kros JM, Chang SM. Diffuse Infiltrating Oligodendroglioma and Astrocytoma. J Clin Oncol. 2017;35(21):2394-401.
15. Gogos AJ, Young JS, Morshed RA, Hervey-Jumper SL, Berger MS. Awake glioma surgery: technical evolution and nuances. J Neurooncol. 2020;147(3):515-24.
16. Brown TJ, Brennan MC, Li M, Church EW, Brandmeir NJ, Rakszawski KL, et al. Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis. JAMA Oncol. 2016;2(11):1460-9.
17. Kubben PL, ter Meulen KJ, Schijns OEMG, ter Laak-Poort MP, van Overbeeke JJ, van Santbrink H. Intraoperative MRI-guided resection of glioblastoma multiforme: A systematic review. Lancet Oncol. 2011;12(11):1062-70.
18. Gessler F, Bernstock JD, Braczynski A, Lescher S, Baumgarten P, Harter PN, et al. Surgery for Glioblastoma in Light of Molecular Markers: Impact of Resection and MGMT Promoter Methylation in Newly Diagnosed IDH-1 Wild-Type Glioblastomas. Neurosurgery. 2018.
19. Li YM, Suki D, Hess K, Sawaya R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? J Neurosurg. 2016;124(4):977-88.
20. Stummer W, Pichlmeier U, Meinel T, Wiestler OD. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. The lancet oncology. 2006.
21. Senft C, Bink A, Franz K, Vatter H, Gasser T, Seifert V. Intraoperative MRI guidance and extent of resection in glioma surgery: a randomised, controlled trial. Lancet Oncol. 2011;12(11):997-1003.
22. 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. J Neurosurg. 2001;95(2):190-8.
23. Kubben PL, Postma AA, Kessels AGH, Van Overbeeke JJ, Van Santbrink H. Intraobserver and interobserver agreement in volumetric assessment of glioblastoma multiforme resection. Neurosurgery. 2010;67(5):1329-34.
24. Unsgaard G, Ommedal S, Muller T, Gronningsaeter A, Nagelhus Hernes TA. Neuronavigation by intraoperative three-dimensional ultrasound: initial experience during brain tumor resection. Neurosurgery. 2002;50(4):804-12; discussion 12.
25. Jenkinson MD, Barone DG, Bryant A, Vale L, Bulbeck H, Lawrie TA, et al. Intraoperative imaging technology to maximise extent of resection for glioma Review. Cochrane Database Syst Rev. 2018;1:CD012788.
26. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, et al. IDH1 and IDH2 mutations in gliomas. N Engl J Med. 2009;360(8):765-73.
27. Lai A, Kharbanda S, Pope WB, Tran A, Solis OE, Peale F, et al. Evidence for sequenced molecular evolution of IDH1 mutant glioblastoma from a distinct cell of origin. J Clin Oncol. 2011;29(34):4482-90.

28. Louis DN, Wesseling P, Aldape K, Brat DJ, Capper D, Cree IA, et al. cIMPACT-NOW Update 6: New Entity and Diagnostic Principle Recommendations of the cIMPACT-Utrecht Meeting on Future CNS Tumor Classification and Grading. Brain Pathol. 2020 Jul;30(4):844-856.
29. Brat DJ, Aldape K, Colman H, Figrarella-Branger D, Fuller GN, Giannini C, et al. cIMPACT-NOW Update 5: Recommended Grading Criteria and Terminologies for IDH-mutant Astrocytomas. Acta Neuropathol . 2020 Mar;139(3):603-608. doi: 10.1007/s00401-020-02127-9. Epub 2020 Jan 29.
30. Eidel O, Burth S, Neumann JO, Kieslich PJ, Sahm F, Jungk C, et al. Tumor Infiltration in Enhancing and Non-Enhancing Parts of Glioblastoma: A Correlation with Histopathology. PLoS One. 2017;12(1):e0169292.
31. Verburg N, Koopman T, Yaqub MM, Hoekstra OS, Lammertsma AA, Barkhof F, et al. Improved detection of diffuse glioma infiltration with imaging combinations: a diagnostic accuracy study. Neuro Oncol. 2020;22(3):412-22.
32. Pessina F, Navarria P, Cozzi L, Ascolese AM, Simonelli M, Santoro A, et al. Maximize surgical resection beyond contrast-enhancing boundaries in newly diagnosed glioblastoma multiforme: is it useful and safe? A single institution retrospective experience. J Neuro-Oncol. 2017;135(1):129-39.
33. Molinaro AM, Hervey-Jumper S, Morshed RA, Young J, Han SJ, Chunduru P, et al. Association of Maximal Extent of Resection of Contrast-Enhanced and Non-Contrast-Enhanced Tumor With Survival Within Molecular Subgroups of Patients With Newly Diagnosed Glioblastoma. JAMA Oncol. 2020;6(4):495-503.
34. Duffau H. Is supratotal resection of glioblastoma in noneloquent areas possible? World Neurosurg. 2014;82(1-2):E101-E3.
35. Aldave G, Tejada S, Pay E, Marigil M, Bejarano B, Idoate MA, et al. Prognostic value of residual fluorescent tissue in glioblastoma patients after gross total resection in 5-aminolevulinic Acid-guided surgery. Neurosurgery. 2013;72(6):915-20; discussion 20-1.
36. Esquenazi Y, Friedman E, Liu Z, Zhu JJ, Hsu S, Tandon N. The Survival Advantage of "Supratotal" Resection of Glioblastoma Using Selective Cortical Mapping and the Subpial Technique. Neurosurgery. 2017;81(2):275-88.
37. Glenn CA, Baker CM, Conner AK, Burks JD, Bonney PA, Briggs RG, et al. An Examination of the Role of Supramaximal Resection of Temporal Lobe Glioblastoma Multiforme. World Neurosurg. 2018;114:e747-e55.
38. Eyüpoglu IY, Hore N, Merkel A, Buslei R, Buchfelder M, Savaskan N. Supra-complete surgery via dual intraoperative visualization approach (DiVA) prolongs patient survival in glioblastoma. Oncotarget. 2016;7(18):25755-68.
39. Hollon TC, Pandian B, Adapa AR, Urias E, Save AV, Khalsa SSS, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. Nat Med. 2020;26(1):52-8.

Samenvatting en conclusie

Dit proefschrift heeft de waarde onderzocht van glioblastoom beeldvorming en resectie in het licht van moleculaire markers. Hier geef ik een samenvatting en bespreek ik de belangrijkste bevindingen in dit proefschrift. Ik bied toekomstperspectieven op onderwerpen die in dit proefschrift aan de orde zijn gekomen en ik sluit af met een conclusie.

In **Deel I** voorspellen we moleculaire markers van gliomen, specifiek IDH-mutatie, 1p/19q codeletie en MGMT promotor methylering status, en cognitieve uitkomst van patiënten met een glioom op basis van preoperatieve MRI scans. In **Deel II** presenteren we de resultaten van een gerandomiseerde, gecontroleerde klinische studie die het nut van intraoperatieve echogeleide chirurgie op de uitgebreidheid van glioblastoom resectie onderzoekt. We beantwoorden de hoofdvraag of intraoperatieve echogeleide chirurgie vaker volledige tumorresectie mogelijk maakt, vergeleken met standaard glioblastoom chirurgie. Op basis van de resultaten geven we een aanbeveling hoe maximale veilige resectie te bereiken. Daarnaast laten we voor het eerst zien middels een concept idee dat augmented reality gebruikt kan worden in de operatiekamer om glioblastoom mee te lokaliseren. In **Deel III** van dit proefschrift onderzoeken we de associatie tussen de mate van resectie van contrast aankleurende en niet-contrast aankleurende tumor en overleving van patiënten met een glioblastoom. We valideren onze bevindingen in een multicenter glioblastoom cohort. Daarnaast onderzochten we het nut van supratotale resectie, dus tumor resectie buiten de grenzen van contrast aankleuring, bij glioblastomen op levensduur van patiënten.

Hoofdbevindingen

Deel I: Preoperatief imaging

In **Hoofdstuk 2** hebben we de topografische verdeling van 436 glioblastomen met vs. zonder MGMT promotor methylering geëvalueerd met behulp van voxel-gewijze MRI analyse in een van de grootste homogene IDH wildtype glioblastoom populaties tot nu toe. Visuele inspectie liet zien dat, vergeleken met MGMT niet-gemethyleerd glioblastomen, gemethyleerde glioblastomen vaker in de buurt van het bifrontale en linker occipitale periventriculaire gebied lagen en minder vaak in de buurt van het rechter occipitale periventriculaire gebied. Deze visuele verschillen konden echter niet worden bevestigd met voxel-gewijze statistische testen. We concluderen daarom dat er geen statistisch verschil was in anatomische lokalisatie tussen IDH wildtype glioblastomen met vs. zonder MGMT promotor methylering. Voorafgaand aan de update van de WHO 2016-classificatie (1) rapporteerden studies tegenstrijdige resultaten; sommige studies toonden dat

glioblastomen met MGMT methylering vaker in de linker temporale kwab zaten, terwijl andere lateralisatie naar de rechterhersenhelft toonden. (2-5) Er waren ook studies die geen verschillen in lokalisatie rapporteerden, in overeenstemming met de bevindingen van onze studie. (6-9) Deze tegenstrijdige resultaten in de literatuur kunnen gedeeltelijk worden verklaard door variaties in de populaties van patiënten met glioblastoom, waarvan er vele werden uitgevoerd voorafgaand aan het tijdperk van de WHO van 2016 classificatie, toen de impact van moleculaire subtypering van glioblastomen op basis van IDH-mutatie status minder relevant geacht werd.

In **Hoofdstuk 3** voorspelden we de moleculaire marker, 1p/19q codeletie, in 284 laaggradige gliomen op basis van preoperatieve MRI-scans door gebruik te maken van machine learning. We hebben ons algoritme extern gevalideerd op een onafhankelijke dataset (n = 129) en we vergeleken de prestatie van ons algoritme met die van klinische experts. We ontdekten dat tumor lokalisatie, tumor heterogeniteit op T2-gewogen MRI scans, samen met leeftijd en geslacht van de patiënt de belangrijkste variabelen waren voor ons algoritme. De nauwkeurigheid in het voorspellen van de 1p/19q codeletie status kwam gemiddeld overeen met die van ervaren oncologische neuroradiologen en presteerde zelfs beter dan oncologische neurochirurgen. Voor zover wij weten, zijn er geen eerdere studies die genetische mutaties van een laaggradige gliomen op MRI voorspellen met behulp van kunstmatige intelligentie die tevens hun algoritme vergeleken met artsen. (10-13) Kennis van het moleculaire profiel van laaggradige glioom is klinisch relevant, aangezien het aanbevolen is om hiermee rekening te houden bij het timen van behandeling met chemo- en of radiotherapie na een operatie. (14) Voordat deze bevindingen klinisch geïmplementeerd kunnen worden in de toekomst, is meer onderzoek nodig om de nauwkeurigheid van algoritmen te verbeteren, maar wij hebben hun potentiaal hier getoond om de besluitvorming in de klinische zorg voor gliomen te ondersteunen.

In **Hoofdstuk 4** hebben we een deep learning algoritme ontwikkeld en deze extern gevalideerd welke de IDH mutatie status, 1p/19q co-deletie status en glioom gradering nauwkeurig kan voorspellen en welke tegelijkertijd automatisch gliomen kan segmenteren, allemaal op basis van preoperatieve MRI scans. Voor patiënten bij wie tumorweefsel niet kan worden verkregen middels een operatie of diagnostische biopsie vanwege hogere leeftijd, slechte neurologische toestand of diepe tumor lokalisatie, kan een nauwkeurige voorspelling van moleculaire markers op preoperatieve MRI scans door algoritmen als alternatief klinisch relevant zijn in de toekomst.

In **Hoofdstuk 5** onderzochten we de associatie tussen cognitie van patiënten met een laaggradige glioom en veranderingen in zenuwbanen (witte stofbanen) met behulp van diffusie tensor MRI en tractografie. We ontdekten dat cognitie in de zin van taal

(herhalen van woorden), inprenting, en aandacht geassocieerd was met veranderde microarchitectuur van de fasciculus arcuata en de fasciculus fronto-occipitalis inferior. Onze resultaten ondersteunen de hypothese dat gliomen zenuwbanen verantwoordelijk voor taal kunnen infiltreren en hun microarchitectuur kunnen verstoren, wat leidt tot verslechtering van de cognitie bij patiënten met glioom. Uitgebreide glioomresectie moet zorgvuldig worden afgewogen tegen het intact laten van dergelijke geïnfiltreerde witte stofkanalen om functie te beschermen, bij voorkeur met een wakkere operatie en (sub)corticale stimulatie. (15)

Deel II: Beeldgeleide glioblastoom operatie

In **Hoofdstuk 6** presenteren we de resultaten van de gerandomiseerde, gecontroleerde, klinische studie die het doel had om het nut van intraoperatieve echogeleide chirurgie te onderzoeken op de mate van resectie bij glioblastomen. De belangrijkste onderzoeksvraag was: maakt intraoperatieve echogeleide chirurgie vaker volledige tumor resectie mogelijk dan standaard glioblastoom chirurgie? We toonden aan dat met intraoperatieve echogeleiding significant vaker (35%) volledige tumorresectie bereikt werd dan met standaard glioblastoom chirurgie (8%). We zagen dat intraoperatieve echogeleiding patiënten niet schaadde in de zin van neurologische uitkomst, functionele status of kwaliteit van leven. Daarom bevelen we intraoperatieve echogeleide chirurgie, in plaats van standaardchirurgie, om volledige en veilige tumor resectie te bereiken.

Volledige resectie van contrast aankleurende tumor is vaker geassocieerd met langere overleving bij patiënten met een glioblastoom. (16) Volledige tumorresectie wordt zowel kwalitatief als kwantitatief gedefinieerd. (17, 18) Sommige studies definieerden volledige tumorresectie als 'geen contrast-aankleurende tumor rest' op een postoperatieve MRI scan, wat een strikte definitie is en kan leiden tot een fout-positief resultaat als gevolg van niet-specifieke contrast aankleuring, zoals ischemie, kleine bloedvaten of atyptische aankleuring van reactief weefsel. (19) Om met dergelijke interpretatie verschillen rekening te houden, definieerden sommige studies ook volledige tumorresectie als 'contrast aankleurende tumor kleiner dan 0,175 cm^3 (20, 21) of als een bepaald resectie-afkappercentage (bijv.> 98%). (22) Het is bekend dat bij het beoordelen van tumor rest bij glioblastomen relatief veel interpretatie verschil zit tussen beoordelaars, wat een zekere mate van subjectiviteit introduceert bij het onderscheiden van contrast aankleurende tumorrest van aspecifieke contrast aankleuring. (23) In ons onderzoek hebben we volledige tumorresectie gedefinieerd als een resectie van > 99% van het contrast aankleurende tumorvolume, waarbij we een contrast aankleurende rest volume van minder dan één procent hebben geaccepteerd om met eventuele interpretatie verschillen rekening te houden.

In twee gerandomiseerde, gecontroleerde, klinische studies is aangetoond dat 5-aminolevulinezuur fluorescentie geleide chirurgie en intraoperatieve MRI geleide chirurgie de mate van glioblastoom resectie verhoogt. (20, 21) Een intra-operatief MRI-systeem is echter duur en verlengt de operatietijd met ongeveer een uur. (21) Dit zou kunnen verklaren waarom intraoperatieve MRI nog steeds niet wereldwijd wordt toegepast als standaard procedure tijdens glioblastoom chirurgie. Een alternatieve, goedkopere en tijdbesparende technologie die wordt gebruikt om terplekke beeldvorming te verkrijgen tijdens hersentumorchirurgie, is intraoperatieve echogeleide chirurgie. (24) Real-time beeldvorming is noodzakelijk, aangezien neuronavigatiesystemen doorgaans gebruik maken van preoperatieve MRI scans. Echter, door verplaatsing van de hersenen tijdens het opereren neemt hun nauwkeurigheid bij het weergeven van de werkelijke situatie tijdens de operatie af. Zoals Jenkinson et al. liet zien in een Cochrane review, was er tot op heden geen gerandomiseerde, gecontroleerde, klinische studie uitgevoerd om de waarde van intraoperatieve echogeleiding te onderzoeken om glioblastoom resectie te maximaliseren. (25) Daarom hebben wij voor het eerst een dergelijk studie uitgevoerd en de resultaten hier in dit proefschrift gepresenteerd. We raden intraoperatieve echogeleiding aan als alternatieve intraoperatieve beeldgeleide technologie, net als intraoperatieve MRI of 5-aminolevulinezuur fluorescentie geleide chirurgie (of een combinatie), om de mate van contrast aankleurende tumorresectie veilig te maximaliseren tijdens glioblastoom chirurgie.

In **Hoofdstuk 7** hebben we als één van de eerste prospectieve klinische pilotstudies, de klinische haalbaarheid en nauwkeurigheid van een dragbare augmented reality bril (Microsoft Hololens) getest voor gebruik in de operatiekamer bij het plannen van glioblastoom operaties. We vergeleken de augmented reality bril met een standaard neuronavigatiesysteem en toonden in een proof of concept met kwantitatieve uitkomstmaten aan dat augmented reality potentieel gebruikt kan worden voor visualisatie en lokalisatie van glioblastomen in het kader van neurochirurgische planning. We vonden dat het handmatig plaatsen van een hologram op de patiënt een beperking was die de nauwkeurigheid van de augmented reality bril verminderde. Voor klinische implementatie moet de nauwkeurigheid van de augmented reality bril daarom nog worden verbeterd in toekomstig onderzoek, bij voorkeur door geautomatiseerde registratie van het hologram op de patiënt en eventueel door integratie van augmented reality met een neuronavigatiesysteem.

Deel III: Mate van resectie en levensduur

In **Hoofdstuk 8** hebben we de associatie tussen contrast aankleurende en niet-contrast aankleurende tumor resectie en overlevingsduur onderzocht in het kader van MGMT promotor methylering bij 326 patiënten met een nieuw gediagnosticeerd IDH wildtype

glioblastoom. Dit onderzoek toonde een verband aan tussen complete contrast aankleurende tumor resectie, ≥30% niet-contrast aankleurende tumor resectie, minimaal contrast aankleurende tumorrest en langere overlevingsduur van patienten. Kaplan Meier grafieken lieten zien dat uitgebreide chirurgie mogelijk gunstiger was voor patiënten met een MGMT gemethyleerd glioblastoom.

Volledige contrast aankleurende tumor resectie was voorafgaand aan dit proefschrift al vaker geassocieerd met langere levensduur van patiënten. (16) Dit proefschrift onderzocht deze associatie echter op basis van twee nieuwe perspectieven.

Ten eerste rapporteerden eerdere studies beperkte gegevens over de moleculaire classificatie van glioblastomen (op basis van IDH-mutatie en MGMT-promotor-methylerings status) aangezien er veel werden uitgevoerd voorafgaand aan de update van de WHO 2016 classificatie, waarbij moleculaire sub classificatie minder een relevant geacht werden. (16) Hoewel IDH-mutatie bij nieuw gediagnosticeerd primair glioblastoom zeldzaam is (<5% en 3,5% in ons cohort), vertegenwoordigen deze tumoren een andere moleculaire glioomtype, afkomstig van een andere precursorlaesie. (26, 27) Daarom is onvolledig of ontbreken van moleculaire gegevens over IDH mutatie en MGMT methylerings status, of het mengen van moleculaire subtypes bij het onderzoeken van het effect van glioblastoom resectie op overleving, ongewenst. In dit proefschrift hebben we de associatie tussen complete contrast aankleurende tumorresectie en langere overleving bevestigd in patiënten met een moleculair gedefinieerd glioblastoom.

In **Hoofdstuk 9** hebben we een internationale, multicenter, observationele studie uitgevoerd met meer dan duizend patiënten met een nieuw gediagnosticeerd IDH-wildtype glioblastoom, waarin we een voorspellingsmodel hebben ontwikkeld en deze extern hebben gevalideerd. Dit model omvatte leeftijd, geslacht, preoperatieve Karnofsky performance status, mate van chirurgische resectie, MGMT-promotor methylatiestatus en adjuvante therapie en toonde na externe validatie dat de overleving van individuele patiënten met een nieuw gediagnosticeerd glioblastoom betrouwbaar kon worden voorspeld. We hebben een een online nomogram gepubliceerd dat bruikbaar kan zijn om gedeelde besluitvorming te ondersteunen.

Ten tweede liet dit proefschrift zien dat glioblastoom resectie geassocieerd is met overlevingsduur, door het nut van niet-contrast aankleurende tumor resectie te laten zien, voorbij de grenzen van contrast aankleuring. Het is bekend dat glioblastoom ver voorbij de contrast aankleurende grenzen, in de omgevende oedemateuze gebieden, infiltreert, wat zichtbaar is als T2-gewogen of FLAIR gewogen hyper intense gebieden op MRI scans.(28) Dit roept de vraag op of complete resectie van contrast aankleurende tumor daadwerkelijk compleet is, of dat resectie voorbij de grenzen van contrast aankleuring, in

de niet-aankleurende delen, uitgebreid moet worden om overlevingsduur van patiënten te verbeteren. De tweede vraag in deze lijn is welk deel van de niet-aankleurende tumor dan gereseceerd zou moeten worden, aangezien verondersteld kan worden dat alleen resectie van tumor geïnfiltreerde weefsel, en niet het oedemateuze hersenweefsel, de hoeveelheid tumor zal verminderen, en dus zal bijdragen aan het overlevingsvoordeel. Fysiologische MRI is een veelbelovend techniek om tumor geïnfiltreerde hersenweefsel te onderscheiden van oedeem in niet contrast aankleurende gebieden.(29) In Hoofdstuk 8 toonden we ook aan dat dat ≥30% niet-aankleurende tumor resectie, voorbij de grenzen van contrast aankleuring, was geassocieerd met een betere overlevingsduur. Andere studies hebben gevonden dat niet-aankleurende tumor resectie percentages van 45%,(30) 53%(19) en 92% (31) was geassocieerd met betere levensduur voor patiënten. Belangrijk is dat maximale resectie van glioblastoom niet koste wat het kost nagestreefd moet worden en over het algemeen wordt in voorgaande studies slechts beperkt over patiënt uitkomsten gerapporteerd. Ook wij hadden onvoldoende gegevens over patiënten uitkomsten in ons onderzoek en daarom adviseren we intraoperatieve beeldvorming en (sub)corticale stimulatie om functionele gebieden te identificeren en daarmee veilig en compleet te opereren. In toekomstig onderzoek moet de veiligheid van uitgebreide glioblastoom resectie verder onderzocht worden.

Binnen de laaggradige glioom chirurgie is het concept van supratotale resectie al langer bekend, maar bij glioblastomen wordt dit nog pas vrij recent onderzocht.(32) **Hoofdstuk 10.1** had tot doel om supratotale resectie bij glioblastomen en de relatie tot levensduur van patiënten te onderzoeken in de literatuur. We lieten in onze systematisch reviewe en meta-analyse gebaseerd op zes studies met 1168 glioblastoom patiënten zien, dat vergeleken met totale resectie, supratotale resectie van glioblastomen geassocieerd was met lager risico op overlijden en langere mediane overleving gedurende follow up. De kwaliteit van het beschikbare bewijs was echter laag door selectiebias en confounding in de geïncludeerde studies.(19, 30, 33-36) Daarbij was er veel variatie in hoe supratotale resectie bij glioblastomen werd gedefinieerd tussen de studies onderling. De veiligheid van supratotale resectie was slecht gerapporteerd, waarbij vaak gegevens misten over neurologische uitkomst, complicaties en kwaliteit van leven van patiënten. Supratotale resectie van glioblastomen kan potentieel bijdrage aan het verbeteren van prognose van patiënten met een glioblastoom, maar meer onderzoek is nodig in de toekomst om de veiligheid ervan te onderzoeken, zoals ook aangegeven in ons eerdere brief (**Hoofdstuk 10.2**).

Toekomstperspectief

We verwachten dat in toekomstig neuro-oncologisch radiogenomics onderzoek de nauwkeurigheid van moleculaire en klinische uitkomst voorspellingen op basis van MRI scans zal toenemen als gevolg van 1) de opkomst van kunstmatige intelligentie 2) meer patiëntgegevens als gevolg van grote multicenter samenwerkingen en consortia en 3) de toevoeging van geavanceerde MRI technieken, waaronder fysiologische MRI scans zoals perfusie en diffusie MRI gecombineerd met PET scan. Fysiologische MRI scans zal niet alleen nuttig zijn voor radiogenomics onderzoek, maar ook voor het veilig maximaliseren van glioom resectie door een nauwkeurigere detectie van de tumor grenzen, beter dan we momenteel kunnen op basis van alleen anatomische MRI-scans. (29) Dit kan ons helpen te begrijpen welk deel van het niet-contrast aankleurende deel verwijderd moet worden vanwege tumor infiltratie en welk deel gespaard kan blijven, omdat er alleen sprake is van oedeem.

We verwachten verder dat nieuwe intraoperatieve beeldvormende technologieën, gecombineerd met kunstmatige intelligentie, nuttig zullen zijn om de mate van tumor resectie veilig te maximaliseren. Een recent onderzoek toonde al het potentiële voordeel aan van intra-operatief beoordelen van tumor grenzen op infiltratie, waarbij binnen enkele minuten met behulp van gestimuleerde Raman-spectroscopie en kunstmatige intelligentie informatie werd verkregen om zo maximale en veilige glioomchirurgie na te streven, die in de toekomst mogelijk vriescoupes zou kunnen vervangen. (37) Deze technologieën, eventueel geïntegreerd met augmented reality-navigatiesystemen, kunnen intraoperatieve navigatie en besluitvorming tijdens chirurgie ondersteunen, bijvoorbeeld door nauwkeurige voorspellingen van algoritmen over weefsel te projecteren om te beslissen of dat laatste deel in de resectieholte resterend tumor of gezond weefsel is, wat vaak moeilijk is te zien zelfs voor een geoefend neurochirurgisch oog. De toekomst van neurochirurgie in die zin zal worden gevormd door nieuwe beeldvormende technologieën en kunstmatige intelligentie, waarin een transitie zal plaatsvinden naar data gedreven chirurgie, waarin nog altijd chirurgische ervaring, intuïtie en anatomische kennis centraal zullen staan in besluitvorming.

Er kan verwacht worden dat dit niet alleen zal bijdragen aan het maximaliseren van de mate van tumor resectie, maar dat ook de veiligheid van hersentumorchirurgie ook zal worden verbeterd door behoud van functioneel hersengebieden. Anatomische beeldgeleide chirurgie zoals intra-operatieve MRI, intraoperatieve echografie en fluorescente geleide chirurgie zullen worden ondersteund door functiegeleide chirurgie zoals een wakkere operatie met elektrische (sub) corticale stimulatie en intraoperatieve fysiologische beeldgeleiding.

Kortom, er kan verwacht worden dat nieuwe beeldvormende technologieën en kunstmatige intelligentie de manier waarop we glioblastomen diagnosticeren, reseceren

en daarmee uiteindelijk de manier waarop we klinische zorg bieden aan patiënten met hersentumoren, zullen verbeteren.

Conclusie

Dit proefschrift draagt bij aan de kennis over de waarde van beeldvorming en resectie van gliomen, in het bijzonder van glioblastomen.

We presenteerden veelbelovende resultaten over moleculaire voorspellingen van gliomen op basis van preoperatieve MRI scans met kunstmatige intelligentie. We toonden aan dat intraoperatieve echogeleide chirurgie vaker tot volledige tumor resectie resulteerde dan met standaard glioblastoom chirurgie, zonder daarbij patiënten te schaden in de zin van neurologische uitkomst, functionele status en kwaliteit van leven. We bevelen intraoperatieve echogeleide chirurgie aan in plaats van standaardchirurgie om een volledig veilige tumorresectie te bereiken. We hebben laten zien dat augmented reality potentieel nuttig was om glioblastomen op de operatiekamer te visualiseren en te lokaliseren in het kader van neurochirurgische planning.

Voorafgaand aan dit proefschrift, was volledige glioblastoom resectie al vaker geassocieerd met een betere prognose voor patiënten met een glioblastoom. We bevestigden deze associatie in moleculair gedefinieerde glioblastomen volgens de geupdate WHO 2016 classificatie. We hebben deze resultaten extern gevalideerd in een multicenter glioblastoom cohort en we voorspelden de overleving van patiënten met een glioblastoom met nomogrammen. Bovendien lieten we zien dat resectie van niet aankleurende tumor deel voorbij de grenzen van contrast aankleuring, dus supratotale resectie, potentieel bijdraagt aan een langere overleving bij patiënten met een glioblastoom. We moedigden verder onderzoek aan naar de veiligheid van het maximaliseren van glioblastoom resectie, vooral buiten de grenzen van contrast aankleuring.

Ten slotte deelden we toekomstperspectieven, waarin we verwachten dat nieuwe beeldvormende technologieën en kunstmatige intelligente zullen bijdragen aan onze kennis, diagnostiek en chirurgische behandeling van patiënten met een hersentumor.

We hopen dat dit proefschrift nuttig zal zijn voor toekomstig hersentumoronderzoek, klinische zorg en voor patiënten met een hersentumor.

Referenties

1. 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 Neuropathol. 2016;131(6):803-20.
2. Ellingson BM, Lai A, Harris RJ, Selfridge JM, Yong WH, Das K, et al. Probabilistic radiographic atlas of glioblastoma phenotypes. Am J Neuroradiol. 2013;34(3):533-40.
3. Ellingson BM, Cloughesy TF, Pope WB, Zaw TM, Phillips H, Lalezari S, et al. Anatomic localization of O6-methylguanine DNA methyltransferase (MGMT) promoter methylated and unmethylated tumors: a radiographic study in 358 de novo human glioblastomas. Neuroimage. 2012;59(2):908-16.
4. Eoli M, Menghi F, Bruzzone MG, De Simone T, Valletta L, Pollo B, et al. Methylation of O6-methylguanine DNA methyltransferase and loss of heterozygosity on 19q and/or 17p are overlapping features of secondary glioblastomas with prolonged survival. Clin Cancer Res. 2007;13(9):2606-13.
5. Wang Y, Fan X, Zhang C, Zhang T, Peng X, Li S, et al. Anatomical specificity of O6-methylguanine DNA methyltransferase protein expression in glioblastomas. J Neurooncol. 2014;120(2):331-7.
6. Han Y, Yan LF, Wang XB, Sun YZ, Zhang X, Liu ZC, et al. Structural and advanced imaging in predicting MGMT promoter methylation of primary glioblastoma: a region of interest based analysis. BMC Cancer. 2018;18(1):215.
7. Carrillo JA, Lai A, Nghiemphu PL, Kim HJ, Phillips HS, Kharbanda S, et al. Relationship between tumor enhancement, edema, IDH1 mutational status, MGMT promoter methylation, and survival in glioblastoma. AJNR Am J Neuroradiol. 2012;33(7):1349-55.
8. Drabycz S, Roldán G, de Robles P, Adler D, McIntyre JB, Magliocco AM, et al. An analysis of image texture, tumor location, and MGMT promoter methylation in glioblastoma using magnetic resonance imaging. Neuroimage. 2010;49(2):1398-405.
9. Roux A, Roca P, Edjlali M, Sato K, Zanello M, Dezamis E, et al. MRI Atlas of IDH Wild-Type Supratentorial Glioblastoma: Probabilistic Maps of Phenotype, Management, and Outcomes. Radiology. 2019;293(3):633-43.
10. Chang P, Grinband J, Weinberg BD, Bardis M, Khy M, Cadena G, et al. Deep-Learning Convolutional Neural Networks Accurately Classify Genetic Mutations in Gliomas. AJNR Am J Neuroradiol. 2018;39(7):1201-7.
11. Han Y, Xie Z, Zang Y, Zhang S, Gu D, Zhou M, et al. Non-invasive genotype prediction of chromosome 1p/19q co-deletion by development and validation of an MRI-based radiomics signature in lower-grade gliomas. J Neurooncol. 2018;140(2):297-306.
12. Zhou H, Chang K, Bai HX, Xiao B, Su C, Bi WL, et al. Machine learning reveals multimodal MRI patterns predictive of isocitrate dehydrogenase and 1p/19q status in diffuse low- and high-grade gliomas. J Neurooncol. 2019;142(2):299-307.
13. Park YW, Han K, Ahn SS, Bae S, Choi YS, Chang JH, et al. Prediction of IDH1-Mutation and 1p/19q-Codeletion Status Using Preoperative MR Imaging Phenotypes in Lower Grade Gliomas. AJNR Am J Neuroradiol. 2018;39(1):37-42.

14. van den Bent MJ, Smits M, Kros JM, Chang SM. Diffuse Infiltrating Oligodendroglioma and Astrocytoma. J Clin Oncol. 2017;35(21):2394-401.
15. Gogos AJ, Young JS, Morshed RA, Hervey-Jumper SL, Berger MS. Awake glioma surgery: technical evolution and nuances. J Neurooncol. 2020;147(3):515-24.
16. Brown TJ, Brennan MC, Li M, Church EW, Brandmeir NJ, Rakszawski KL, et al. Association of the Extent of Resection With Survival in Glioblastoma: A Systematic Review and Meta-analysis. JAMA Oncol. 2016;2(11):1460-9.
17. Kubben PL, ter Meulen KJ, Schijns OEMG, ter Laak-Poort MP, van Overbeeke JJ, van Santbrink H. Intraoperative MRI-guided resection of glioblastoma multiforme: A systematic review. Lancet Oncol. 2011;12(11):1062-70.
18. Gessler F, Bernstock JD, Braczynski A, Lescher S, Baumgarten P, Harter PN, et al. Surgery for Glioblastoma in Light of Molecular Markers: Impact of Resection and MGMT Promoter Methylation in Newly Diagnosed IDH-1 Wild-Type Glioblastomas. Neurosurgery. 2018.
19. Li YM, Suki D, Hess K, Sawaya R. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: Can we do better than gross-total resection? J Neurosurg. 2016;124(4):977-88.
20. Stummer W, Pichlmeier U, Meinel T, Wiestler OD. Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase III trial. The lancet oncology. 2006.
21. Senft C, Bink A, Franz K, Vatter H, Gasser T, Seifert V. Intraoperative MRI guidance and extent of resection in glioma surgery: a randomised, controlled trial. Lancet Oncol. 2011;12(11):997-1003.
22. 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. J Neurosurg. 2001;95(2):190-8.
23. Kubben PL, Postma AA, Kessels AGH, Van Overbeeke JJ, Van Santbrink H. Intraobserver and interobserver agreement in volumetric assessment of glioblastoma multiforme resection. Neurosurgery. 2010;67(5):1329-34.
24. Unsgaard G, Ommedal S, Muller T, Gronningsaeter A, Nagelhus Hernes TA. Neuronavigation by intraoperative three-dimensional ultrasound: initial experience during brain tumor resection. Neurosurgery. 2002;50(4):804-12; discussion 12.
25. Jenkinson MD, Barone DG, Bryant A, Vale L, Bulbeck H, Lawrie TA, et al. Intraoperative imaging technology to maximise extent of resection for glioma Review. Cochrane Database Syst Rev. 2018;1:CD012788.
26. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, et al. IDH1 and IDH2 mutations in gliomas. N Engl J Med. 2009;360(8):765-73.
27. Lai A, Kharbanda S, Pope WB, Tran A, Solis OE, Peale F, et al. Evidence for sequenced molecular evolution of IDH1 mutant glioblastoma from a distinct cell of origin. J Clin Oncol. 2011;29(34):4482-90.

28. Louis DN, Wesseling P, Aldape K, Brat DJ, Capper D, Cree IA, et al. cIMPACT-NOW Update 6: New Entity and Diagnostic Principle Recommendations of the cIMPACT-Utrecht Meeting on Future CNS Tumor Classification and Grading. Brain Pathol. 2020 Jul;30(4):844-856.
29. Brat DJ, Aldape K, Colman H, Figrarella-Branger D, Fuller GN, Giannini C, et al. cIMPACT-NOW Update 5: Recommended Grading Criteria and Terminologies for IDH-mutant Astrocytomas. Acta Neuropathol . 2020 Mar;139(3):603-608. doi: 10.1007/s00401-020-02127-9. Epub 2020 Jan 29.
30. Eidel O, Burth S, Neumann JO, Kieslich PJ, Sahm F, Jungk C, et al. Tumor Infiltration in Enhancing and Non-Enhancing Parts of Glioblastoma: A Correlation with Histopathology. PLoS One. 2017;12(1):e0169292.
31. Verburg N, Koopman T, Yaqub MM, Hoekstra OS, Lammertsma AA, Barkhof F, et al. Improved detection of diffuse glioma infiltration with imaging combinations: a diagnostic accuracy study. Neuro Oncol. 2020;22(3):412-22.
32. Pessina F, Navarria P, Cozzi L, Ascolese AM, Simonelli M, Santoro A, et al. Maximize surgical resection beyond contrast-enhancing boundaries in newly diagnosed glioblastoma multiforme: is it useful and safe? A single institution retrospective experience. J Neuro-Oncol. 2017;135(1):129-39.
33. Molinaro AM, Hervey-Jumper S, Morshed RA, Young J, Han SJ, Chunduru P, et al. Association of Maximal Extent of Resection of Contrast-Enhanced and Non-Contrast-Enhanced Tumor With Survival Within Molecular Subgroups of Patients With Newly Diagnosed Glioblastoma. JAMA Oncol. 2020;6(4):495-503.
34. Duffau H. Is supratotal resection of glioblastoma in noneloquent areas possible? World Neurosurg. 2014;82(1-2):E101-E3.
35. Aldave G, Tejada S, Pay E, Marigil M, Bejarano B, Idoate MA, et al. Prognostic value of residual fluorescent tissue in glioblastoma patients after gross total resection in 5-aminolevulinic Acid-guided surgery. Neurosurgery. 2013;72(6):915-20; discussion 20-1.
36. Esquenazi Y, Friedman E, Liu Z, Zhu JJ, Hsu S, Tandon N. The Survival Advantage of "Supratotal" Resection of Glioblastoma Using Selective Cortical Mapping and the Subpial Technique. Neurosurgery. 2017;81(2):275-88.
37. Glenn CA, Baker CM, Conner AK, Burks JD, Bonney PA, Briggs RG, et al. An Examination of the Role of Supramaximal Resection of Temporal Lobe Glioblastoma Multiforme. World Neurosurg. 2018;114:e747-e55.
38. Eyüpoglu IY, Hore N, Merkel A, Buslei R, Buchfelder M, Savaskan N. Supra-complete surgery via dual intraoperative visualization approach (DiVA) prolongs patient survival in glioblastoma. Oncotarget. 2016;7(18):25755-68.
39. Hollon TC, Pandian B, Adapa AR, Urias E, Save AV, Khalsa SSS, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. Nat Med. 2020;26(1):52-8.

Chapter 12

Acknowledgement
PhD portfolio
About the author

Acknowledgement · Dankwoord

De volgende personen hebben mij geholpen bij het tot stand brengen van dit proefschrift en hebben tevens een belangrijke bijdrage geleverd in mijn ontwikkeling als mens, arts en wetenschapper. Graag wil ik ze hier persoonlijk bedanken.

Prof. dr. Marion Smits
Beste Marion, toen de onderzoeksgeld die ik bij de neurochirurgie had binnengehaald onvoldoende bleek te zijn voor een volledige promotietraject, heb jij mij de kans aangeboden om toch promotieonderzoek te doen. Daarnaast heb je het mij gegund om de opleiding epidemiologie te volgen en je vond het goed dat ik filosofie ging studeren naast mijn promotietraject. Ik heb alle vrijheid van je gekregen om zelf sturing te geven aan mijn promotieonderzoek. Dank voor alle ruimte en alle mogelijkheden die ik van je heb gekregen. Je bent niet alleen een scherpe wetenschapper, maar ook een inspirerende mentor met een sterk moreel kompas. Ik ben dankbaar dat ik je heb leren kennen, dat ik van je heb kunnen leren en dat ik met je heb mogen samenwerken. Ik hoop dat we in de toekomst zullen blijven samenwerken.

Prof. dr. Martin van den Bent
Beste Martin, ik heb de afgelopen jaren geluk gehad dat ik kon leren van jouw expertise en ervaring als internationale glioom onderzoeker en neuro-oncoloog. Bedankt voor de adviezen die ik van je heb gehad over zowel mijn promotie onderzoek als daarbuiten. Ik heb genoten van onze telkens terugkomende discussie over waar het nou daadwerkelijk allemaal om gaat; mate van resectie of tumor residu?

Dr. Arnaud Vincent
Beste Arnaud, jij hebt mij in 2010 als 2e jaars geneeskundestudent geënthousiasmeerd voor glioomonderzoek en mij de afgelopen 10 jaar in mijn onderzoek gesteund. Ook heb je mij toen als co-assistent geholpen om de ultrasound trial op te zetten en daar een grant voor binnen te halen. Daarnaast heb je mij ook gesteund in mijn idee om augmented reality te gebruiken voor glioblastoom operaties. Dit alles heeft tot een belangrijk deel van dit proefschrift geresulteerd, bedankt hiervoor.

Leden van de leescommissie
Beste Prof. dr. P.C. de Witt Hamer, Prof. dr. M.W. Vernooij, Prof. dr. M.J.B. Taphoorn; het is een eer dat jullie met kennis en ervaring binnen jullie eigen vakgebied in de neurochirurgie, neuroradiologie en neuro-oncologie mijn proefschrift hebben beoordeeld, bedankt dat jullie hebben willen deelnemen aan mijn leescommissie.

Sebastian van der Voort
Beste Sebastian, wij zijn tegelijkertijd begonnen met onze promotie trajecten, jij aan de technische kant en ik aan de klinische kant. Dankzij jou heb ik de afgelopen jaren veel geleerd over kunstmatige intelligentie. Wat ik als simpele arts handmatig in dagen deed kon jij veel sneller met je algoritme. We zijn in het begin vooral bezig geweest met het verzamelen van klinische, radiologische en genetische data van honderden gliomen. Uiteindelijk zijn daar mooie artikelen uit voortgekomen en hebben we een basis gelegd voor glioom onderzoekers die na ons zullen komen, daar ben ik trots op. Bedankt voor al je hulp en de prettige samenwerking.

NA-25 kamerleden
Kars, Dianne, Rebecca, Noor, Tiara, we hebben veel te laat een koffieapparaat aangeschaft, maar er daarna wel volop van genoten. Ook heb ik genoten van onze diep filosofische discussies, die ik zeker zal missen.Dank voor de mooie tijden op de 25ste verdieping en succes met het afronden van jullie promotieonderzoeken.

Onderzoeksgroep neuro-oncologie
Beste Sebastian, Karin, Wouter, Renske, Esther, Stephan, Fatemah and Yulun, de neuro-oncologie groep heeft zich in de afgelopen jaren ontwikkeld tot een ware brug tussen klinisch en technisch hersentumor onderzoek. Ik hoop dat de technische ontwikkelingen met nieuwe MRI technieken en kunstmatige intelligentie daadwerkelijk zullen bijdrage aan betere zorg voor patiënten met een hersentumor. De toekomst is wat dat betreft spannend en hoopgevend. Bedankt voor de prettige samenwerking en alle hulp de afgelopen jaren.

Afdeling moleculaire pathologie
Beste Peggy, dr. Erik Jan Dubbink, prof. J.M. Kros en medewerkers van het lab, dankzij jullie is het gelukt om in een vrij korte tijd honderden glioblastomen moleculair te analyseren. Een belangrijk deel van dit proefschrift is gebaseerd op deze data, dank daarvoor!

Neurochirurgen Erasmus MC
Beste neurochirurgen, bedankt voor jullie hulp tijdens mijn klinische werkzaamheden op de afdeling en ook voor alle inzet voor de ultrasound trial.

Poli & OK-planning neurochirurgie
Beste Marit en Dianne, vooral in het oude poli gebouw heb ik soms uren moeten wachten om studie patiënten te kunnen includeren voor de ultrasound trial. Jullie hulp daarbij is groot geweest, dank. Beste Monique en Anneke, zonder jullie hulp zou het plannen van operaties voor studie patiënten onmogelijk zijn.

Trialbureau Radiologie en MRI planning
Beste mensen van de trialbureau en MRI planning, in het bijzonder Laurens en Renée, dankzij jullie assistentie, van de METC aanvraag tot het inplannen van MRIs, kon de iGENE studie gerealiseerd worden, dank!

iGENE Haaglanden MC & Utrecht MC
Beste Rishi Nandoe Tewarie, Geert Lycklama en Claudine Nogarede van het HMC, Jan Willem Dankbaar, Ilse de Vos en de trialbureau van het UMCU, ondanks alle trage bureaucratie is het gelukt om met een deel van de data een aantal mooie artikelen te publiceren. Bedankt voor alle hulp.

Adem Dereci
Beste Adem, ik ben blij en vereerd dat je mijn paranimf bent. Je bent niet alleen mij studiegenoot, maar ook mijn familievriend waar ik altijd ogendicht op kan vertrouwen. Je bent overdag een clinicus in hart en nieren en in de avonden ben je altijd wel een case report aan het schrijven. Hoeveel heb je er nou nu?

Schoonfamilie
Beste schoonfamilie, bedankt voor de interesse die jullie altijd in mijn onderzoek hebben getoond, voor het luisterend oor en voor al jullie steun tijdens mijn promotietraject.

Anne, baba, kardeşim Muhammed
Bir ogul ve bir abi olarak istedigim kadar sizin ile vakit geciremedim. Kendimi bildim bileli ya kutuphanede yada hasthanede calisiyorum. Bunun ileridede cok degisecegini zannetmiyorum, bunu umitte etmiyorum. Anlayisiniz ve yardiminiz icin tesekkurler. Bu kitapcigi size armagan ediyorum.

Anne, baba, universitelerde tib olsun, felsefe olsun, bircok hocadan ders aldim. Hicbiri sizin bana verdiginiz Islami ve ahlaki degerlerden fazla ruhumda ve karakterimde etki birakmadi. Yaptigim ve ogrendigim herseyi bu degerler uzerine bina ettim. Bu yuzden bu kitapta, basarilarimda ve insallah basaracaklarimda sizin etkiniz cok buyuk. Ilime katki saglayabilme ve insanlara faydali olabilme umidi ile calismam sizi umarim mutlu eder. Sara yi buyutmede cok buyuk payiniz var, Merve ile bana cok destek oluyorsunuz. Allah sizden razi olsun. Sizi seviyorum.

Merve Ciftci-Incekara,
Merveciğim, jouw steun en liefde speelt een grote rol in hoe ik mijn ambities in het leven hoopvol nastreef en realiseer. Je bent voor mij een onuitputtelijke bron van positieve energie en liefde. Ik ben dankbaar dat je mijn paranimf bent. Succes met je opleiding tot advocaat, jij kan het! Sara en ik zijn trots op jou en wij houden van jou.

Sara Incekara

Saraciğim, sen annenin ve benim mutluluk kaynagimizsin. Umarim hayatin boyunca iman, ahlak ve ilim yolunda ilerlersin ve umarim bu kitapcik senin insanliga faydali bir hayat yasaman icin bir motivasyon ve umit kaynağı olur. Seni seviyorum.

PhD Portfolio

Name PhD Student:	Fatih Incekara
Research school:	Netherlands Institute for Health Sciences
Erasmus MC Departments:	Radiology and Neurosurgery
PhD period:	August 2016 - August 2020
Promotors:	Prof. dr. Marion Smits and Prof. dr. Martin van den Bent
Copromotor:	dr. A.J.P.E. Vincent

PhD Training	Year	ECTS
General courses		
Clinical Epidemiology, Master of Science in Health Sciences (NIHES), courses:	2017-2019	70
Biostatistical Methods I: Basic Principles Biostatistical Methods II: Classical Regression Models Clinical Trials Clinical Epidemiology Fundamentals of Medical Decision-Making Principles of Research in Medicine and Epidemiology Clinical Translation of Epidemiology Principles in Causal Inference Methods of Public Health Research The Practice of Epidemiologic Analysis Conceptual Foundation of Epidemiologic Study Design	Logistic Regression Topics in Meta-analysis Advanced Topics in Clinical Trials Advanced Analysis of Prognosis Studies Advances in Clinical Epidemiology Causal Mediation Analysis Health Economics Markers and Prediction Research Biomedical English Writing Course Study Design	
Philosophy (of medicine), Bachelor of Arts (BA), courses:	2019-2020	90
Philosophy of Science I Philosophy of Science II Philosophy of Science III Descartes and Spinoza Thought Experiments The Early Enlightenment The High Enlightenment I: J.J. Rousseau The High Enlightenment II: I. Kant Aesthetics Thesis: Cartesian Consciousness, an interpretation beyond mind-body dualism	Biomedisch Ethiek Oefeningen in Perspectivisme Human Enhancement Moral Philosophy The Quest for Man I The Quest for Man II: S. Kierkegaard Eco philosophy Social and Political Philosophy Contemporary Challenges Critique	
Research integrity	2016	0.3
Basiscursus Regelgeving en Organisatie voor Klinisch Onderzoekers (BROK)	2016	1.5
Basic course in applied MR techniques (ESMRMB)	2016	1
Diffusion MRI Course (ISMRM, Lisbon)	2016	1.5
Clinical fMRI & DTI Course (ESMRMB, London)	2014	1.5

Conferences and presentations	Year	ECTS
European Association of Neurosurgical Societies (EANS)	2020	1.5
Oral presentation: Intraoperative ultrasound guided surgery and the extent of glioblastoma resection: a randomized, controlled trial. Oral presentation: The association between the extent of glioblastoma resection and survival in light of MGMT promoter methylation in 326 patients with newly diagnosed IDH wildtype glioblastoma		
European Association of Neurosurgical Societies (EANS), Venice Dutch-Belgian Neurosurgical Societies, scientific meeting	2017	1.5
Posters (2x): Clinical feasibility and accuracy of a mixed reality device in neurosurgery		
Dutch Neurosurgical Society (NVvN), wintermeeting European Society for Neuroradiology, Rotterdam	2018-2019	1
Huisartsen Congres "Ontwikkelingen in de Geneeskunde", Rotterdam		
Oral presentations (3x): Clinical feasibility and accuracy of a mixed reality device in neurosurgery		
Dutch Neurosurgical Society (NVvN), wintermeeting	2019	0.3
Oral presentation: Artificial intelligence in oncological neurosurgery		

Teaching activities		
Supervising three medical students for master thesis	2017-2019	2
Fiber tractography and tumor segmentation in neurosurgery	2016	0.3

Grants & Awards	
Best abstract award for oral presentation during EANS 2020: Intraoperative ultrasound guided surgery and the extent of glioblastoma resection: a randomized, controlled trial	2020
Research grant awarded (40k) for the ultrasound trial, Stichting Coolsingel	2016

About the author

Fatih Incekara was born on February 14th, 1990 in Rotterdam. He studied medicine at the Erasmus Medical Center (2009-2015). As a second-year medical student, he started neurosurgical research in image guided neuro-oncological surgery, supervised by dr. Arnaud Vincent. In 2013, he spent 6 months as a visiting researcher at the department of neurosurgery, Brigham and Women's Hospital, Harvard Medical School in Boston, supervised by Prof. dr. Alexandra Golby. Fatih worked as a neurosurgical resident (ANIOS) at the Erasmus Medical Center for 9 months. During this time, he was awarded a research grant for the ultrasound trial, which eventually resulted in a part of this thesis.

In 2016, he started his PhD training on imaging and resection of glioma at the departments of Radiology and Neurosurgery, supervised by Prof. Marion Smits, Prof. Martin van den Bent and dr. Arnaud Vincent. During his PhD training, Fatih completed programmes in clinical epidemiology (Master of Science, NIHES) and philosophy (Bachelor of Arts, Erasmus University Rotterdam).

Fatih is married to Merve, together they have a daughter named Sara.

my Lord, increase my knowledge

Quran 20:114

Medice, cura te ipsum