Machine Learning in Image-Based Glioma Grading

Subjects: Neuroimaging

Contributor: Sara Merkaj , Ryan Bahar , , Ichiro Ikuta , Seyedmehdi Payabvash , Mariam S. Aboian

Technological innovation has enabled the development of machine learning (ML) tools that aim to improve the practice of radiologists. In the last decade, ML applications to neuro-oncology have expanded significantly, with the pre-operative prediction of glioma grade using medical imaging as a specific area of interest.

clinical implementation. Keywords: artificial intelligence   glioma   machine learning

deep learning

1. Introduction

1.1. Artificial Intelligence, Machine Learning, and Radiomics

Innovations in computation and imaging have rapidly enhanced the potential for artificial intelligence (AI) to impact diagnostic neuroradiology. Emerging areas of implementation include AI in stroke (e.g., early diagnosis, detection of large vessel occlusion, and outcome prediction) [1], AI in spine (fracture detection, and vertebrae segmentation) and detection of intracranial aneurysms and hemorrhage [2], among other disciplines. Machine learning (ML) and its subfield, deep learning (DL), are branches of AI that have received particular attention. ML algorithms, including DL, decipher patterns in input data and independently learn to make predictions [3]. The advent of radiomics— which mines data from images by transforming them into features quantifying tumor phenotypes—has fueled the application of ML methods to imaging, including radiomics-based ML analysis of brain tumors [4][5][6].

1.2. Machine Learning Applications in Neuro-Oncology

As the most common primary brain tumors, gliomas constitute a major focus of ML applications to neuro-oncology [7][8]. Prominent domains of glioma ML research include the image-based classification of tumor grade and prediction of molecular and genetic characteristics. Genetic information is not only instrumental to tumor diagnosis in the 2021 World Health Organization classification, but also significantly affects survival and underpins sensitivity to therapeutic interventions [9][10]. ML-based models for predicting tumor genotype can therefore guide earlier diagnosis, estimation of prognosis, and treatment-related decision-making [11][12]. Other significant areas of glioma ML research relevant to neuroradiologists include automated tumor segmentation on MRI, detection and prediction of tumor progression, differentiation of pseudo-progression from true progression, glioma survival prediction and treatment response, distinction of gliomas from other tumors and non-neoplastic lesions, heterogeneity

assessment based on imaging features, and clinical incorporation of volumetrics [13][14][15]. Furthermore, ML tools may optimize neuroradiology workflow by expediting the time to read studies from image review to report generation [16]. As an image interpretation support tool, ML importantly may improve diagnostic performance [17][18].

2. Algorithms for Glioma Grade Classification

The most common high-performing ML classifiers for glioma grading in the literature are SVM and CNN [13]. SVM is a classical ML algorithm that represents objects as points in an n-dimensional space, with features serving as coordinates. SVMs use a hyperplane, or an n-1 dimensional subspace, to divide the space into disconnected areas [19]. These distinct areas represent the different classes that the model can classify. Unlike CNNs, SVMs require hand-engineered features, such as from radiomics, to serve as inputs. This requirement may be advantageous for veteran diagnostic imagers, whose knowledge of brain tumor appearance may enhance feature design and selection. Hand-engineered features also can undergo feature reduction to mitigate the risks of overfitting, and prior works demonstrate better performance for glioma grading models using a smaller number of quantitative features [20]. However, hand-engineered features are limited since they cannot be adjusted during model training, and it is uncertain if they are optimal features for classification. Moreover, hand-engineered features may not generalize well beyond the training set and should be tested extensively prior to usage [21][22].

CNNs are a form of deep learning based on image convolution. Images are the direct inputs to the neural network, rather than the manually engineered features of classical ML. Numerous interconnected layers each compute feature representations and pass them on to subsequent layers [22][23]. Near the network output, features are flattened into a vector that performs the classification task. CNNs appeared for glioma grading in 2018 and have risen quickly in prevalence while exhibiting excellent predictive accuracies [24][25][26][27]. To a greater extent than classical ML, they are suited for working with large amounts of data, and their architecture can be modified to optimize efficiency and performance [25]. Disadvantages include the opaque "black box" nature of deep learning and associated difficulty with interpreting model parameters, along with problems that variably apply to classical ML as well (e.g., high amount of time and data required for training, hardware costs, and necessary user expertise) [28][29].

3. Challenges in Image-Based ML Glioma Grading

3.1. Data Sources

Since 2011, a significant number of ML glioma grade prediction studies have used open-source multi-center datasets to develop their models. BraTS [30] and TCIA [31] are two prominent public datasets that contain multi-modal MRI images of high- and low-grade gliomas and patient demographics. BraTS was first made available in 2012, with the 2021 dataset containing 8000 multi-institutional, multi-parametric MR images of gliomas [30]. TCIA first went online in 2011 and contains MR images of gliomas collected across 28 institutions [31]. These datasets were developed with the aim of providing a unified multi-center resource for glioma research. A variety of predictive models have been trained and tested on these large datasets since their 2011 release [32]. Despite their value as

public datasets for model development, several limitations should be considered. Images are collected across multiple institutions with variable protocols and image quality. Co-registration and imaging pre-processing integrate these images into a single system.

3.2. External Validation

Publications have reported predictive models for glioma grading throughout the last 20 years with the majority relying on internal validation techniques, of which cross-validation is the most popular. While internal validation is a well-established method for measuring how well a model will perform on new cases from the initial dataset, additional evaluation on a separate dataset (i.e., external validation) is critical to demonstrate model generalizability. External validation mitigates site bias (differences amongst centers in protocols, techniques, scanner variability, level of experience, etc.) and sampling/selection bias (performance only applicable to the specific training set population/demographics) [33]. Not controlling for these two major biases undermines model generalizability, yet few publications externally validate their models [13]. Therefore, normalizing external validation is a crucial step in developing glioma grade prediction models that are suitable for clinical implementation.

3.3. Glioma Grade Classification Systems

The classification of glioma subtypes into high- and low-grade gliomas is continuously evolving. In 2016, an integrated histological–molecular classification replaced the previous purely histopathological classification [34]. In 2021, the Consortium to Inform Molecular and Practical Approaches to CNS Tumor Taxonomy (cIMPACT NOW) once more accentuated the diagnostic value of molecular markers, such as the isocitrate dehydrogenase mutation, for glioma classification [35]. As a result of the evolving glioma classification system, definitions for high- and low-grade gliomas vary across ML glioma grade prediction studies and publication years. This reduces the comparability of models themselves and grade-labeled datasets used for model development.

Current and future ML methods must keep abreast of the rapid progress in tissue based integrated diagnostics in order to contribute to and make an impact on the clinical care of glioma patients (**Figure 1**).
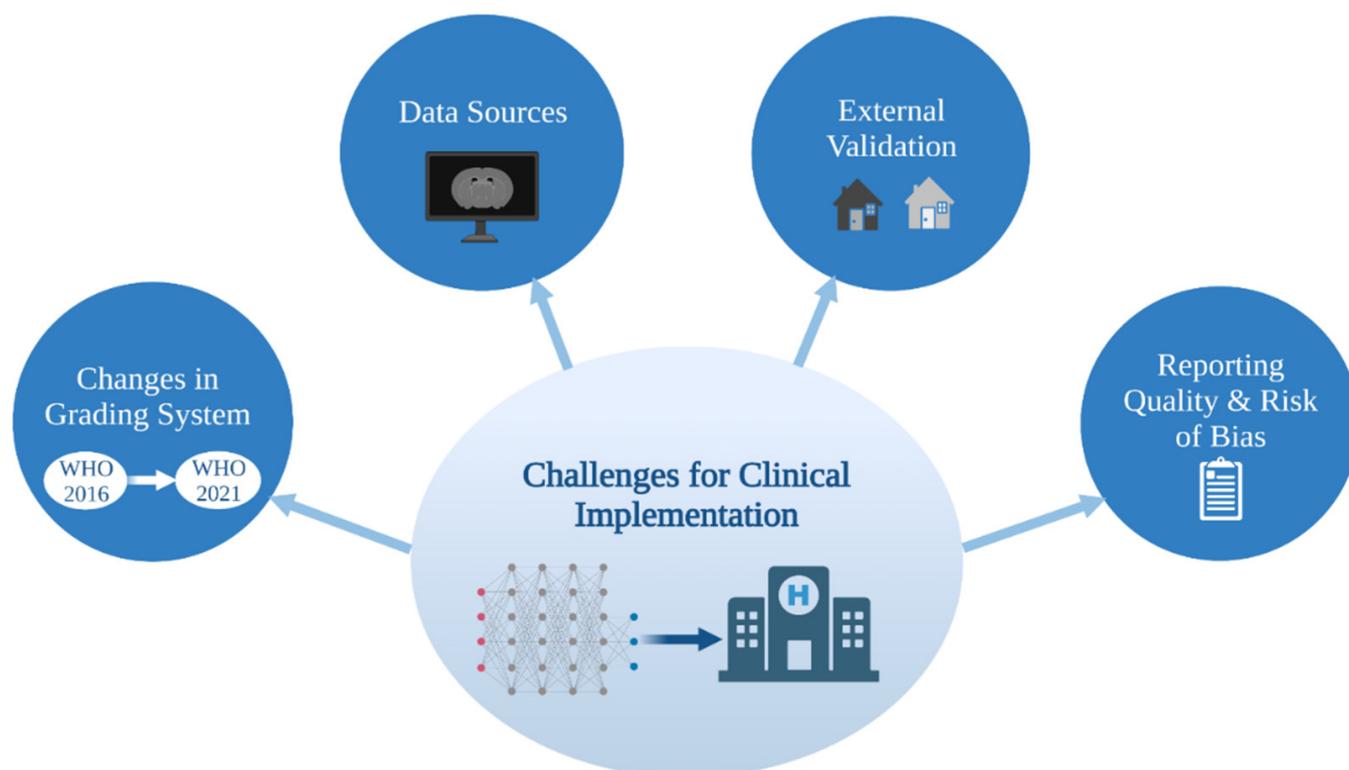
Figure 1. Challenges for clinical implementation of ML glioma grade prediction models. ML = machine learning. WHO = World Health Organization.

3.4. Reporting Quality and Risk of Bias

3.4.1. Overview of Current Guidelines and Tools for Assessment

Clear and thorough reporting enables more complete understanding by the reader and unambiguous assessment of study generalizability, quality, and reproducibility, encouraging future researchers to replicate and use models in clinical contexts. Several instruments have been designed to improve the reporting quality (defined here as the transparency and thoroughness with which authors share key details of their study to enable proper interpretation and evaluation) of studies developing models. The Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) Statement was created in 2015 as a set of recommendations for studies developing, validating, or updating diagnostic or prognostic models [36]. The TRIPOD Statement is a checklist of 22 items considered essential for transparent reporting of a prediction model study. In 2017, with a concurrent rise in radiomics-based model studies, the radiomics quality score (RQS) emerged [37]. RQS is an adaptation of the TRIPOD approach geared toward a radiomics-specific context. The tool has been used throughout the literature for evaluating the methodological quality of radiomics studies, including applications to medical imaging [38]. Radiomics-based approaches for interpreting medical images have evolved to encompass the AI techniques of classical ML and, most recently, deep learning models. Most recently, in recognition of the growing need for an evaluation tool specific to AI applications in medical imaging, the Checklist for AI in Medical Imaging (CLAIM) was published in 2020 [39]. The 42 elements of CLAIM aim to be a best practice guide for authors

presenting their research on applications of AI in medical imaging, ranging from classification and image reconstruction to text analysis and workflow optimization. Other tools—the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool [40] and Prediction model Risk Of Bias ASsessment Tool (PROBAST) [41]—importantly evaluate the risk of bias in studies based on what is reported about their models. Bias relates to systematic limitations or flaws in study design, methods, execution, or analysis that distort estimates of model performance [40]. High risk of bias discourages adaptation of the reported model outside of its original research context, and, at a systemic level, undermines model reproducibility and translation into clinical practice.

3.4.2. Reporting Quality and Risk of Bias in Image-Based Glioma Grade Prediction

Assessments of ML-based prediction model studies have demonstrated that risk of bias is high and reporting quality is inadequate. In their systematic review of prediction models developed using supervised ML techniques, Navarro et al. found that the high risk of study bias, as assessed using PROBAST, stems from small study size, poor handling of missing data, and failure to deal with model overfitting [42].

3.4.3. Future of Reporting Guidelines and Risk of Bias Tools for ML Studies

Efforts by authors to refine how they report their studies depend upon existing reporting guidelines. In their systematic review, Yao et al. identified substantial limitations to neuroradiology deep learning reporting standardization and reproducibility [43]. They recommended that future researchers propose a reporting framework specific to deep learning studies. This call for an AI-targeted framework parallels contemporary movements to produce AI extensions of established reporting guidelines.

4. Future Directions

ML models present an attractive solution towards overcoming current barriers and accelerating the transition to patient-tailored treatments and precision medicine. Novel algorithms combine information derived from multimodal imaging to molecular markers and clinical information, with the aim of bringing personalized predictions on a patient level into routine clinical care. Relatedly, multi-omic approaches that integrate a variety of advanced techniques such as proteomics, transcriptomics, epigenomics, etc., are increasingly gaining importance in understanding cancer biology and will play a key role in the facilitation of precision medicine [44][45]. The growing presence of ML models in research settings is indisputable, yet several strategies should be considered to facilitate clinical implementation: PACS-based image annotation tools, data-sharing and federated learning, ML fairness, ML transparency, and FDA clearance and real-world use (**Figure 2**).
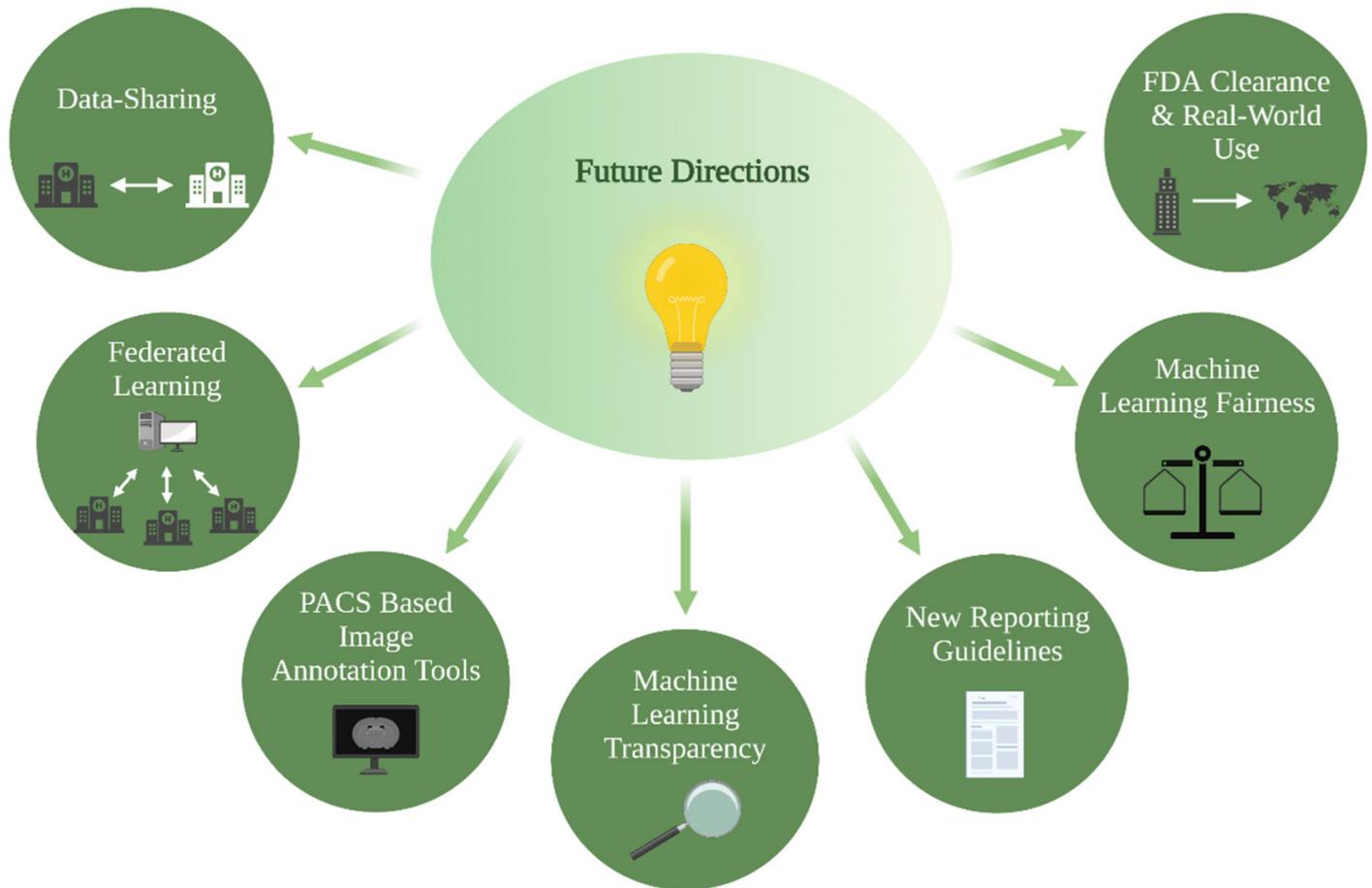
Figure 2. Future directions for clinical implementation of ML glioma grade prediction models, ML = machine learning.

References

1. Yeo, M.; Kok, H.K.; Kutaiba, N.; Maingard, J.; Thijs, V.; Tahayori, B.; Russell, J.; Jhamb, A.; Chandra, R.V.; Brooks, M.; et al. Artificial intelligence in clinical decision support and outcome prediction–Applications in stroke. J. Med. Imaging Radiat. Oncol. 2021, 65, 518–528.

2. Kaka, H.; Zhang, E.; Khan, N. Artificial Intelligence and Deep Learning in Neuroradiology: Exploring the New Frontier. Can. Assoc. Radiol. J. 2021, 72, 35–44.

3. Sidey-Gibbons, J.A.M.; Sidey-Gibbons, C.J. Machine learning in medicine: A practical introduction. BMC Med. Res. Methodol. 2019, 19, 64.

4. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016, 278, 563–577.

5. Aerts, H.J.W.L.; Velazquez, E.R.; Leijenaar, R.T.H.; Parmar, C.; Grossmann, P.; Carvalho, S.; Bussink, J.; Monshouwer, R.; Haibe-Kains, B.; Rietveld, D.; et al. Decoding tumour phenotype by

noninvasive imaging using a quantitative radiomics approach. Nat. Commun. 2014, 5, 4006.

6. Giger, M.L. Machine Learning in Medical Imaging. J. Am. Coll. Radiol. 2018, 15, 512–520.

7. Ostrom, Q.T.; Cioffi, G.; Gittleman, H.; Patil, N.; Waite, K.; Kruchko, C.; Barnholtz-Sloan, J.S. CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012–2016. Neuro-Oncology 2019, 21 Suppl. S5, v1–v100.

8. Abdel Razek, A.A.K.; Alksas, A.; Shehata, M.; AbdelKhalek, A.; Abdel Baky, K.; El-Baz, A.; Helmy, E. Clinical applications of artificial intelligence and radiomics in neuro-oncology imaging. Insights Imaging 2021, 12, 152.

9. Thon, N.; Tonn, J.-C.; Kreth, F.-W. The surgical perspective in precision treatment of diffuse gliomas. OncoTargets Ther. 2019, 12, 1497–1508.

10. Hu, L.S.; Hawkins-Daarud, A.; Wang, L.; Li, J.; Swanson, K.R. Imaging of intratumoral heterogeneity in high-grade glioma. Cancer Lett. 2020, 477, 97–106.

11. Seow, P.; Wong, J.H.D.; Annuar, A.A.; Mahajan, A.; Abdullah, N.A.; Ramli, N. Quantitative magnetic resonance imaging and radiogenomic biomarkers for glioma characterisation: A systematic review. Br. J. Radiol. 2018, 91, 20170930.

12. Lambin, P.; Rios-Velazquez, E.; Leijenaar, R.; Carvalho, S.; van Stiphout, R.G.P.M.; Granton, P.; Zegers, C.M.L.; Gillies, R.; Boellard, R.; Dekker, A.; et al. Radiomics: Extracting more information from medical images using advanced feature analysis. Eur. J. Cancer 2012, 48, 441–446.

13. Buchlak, Q.D.; Esmaili, N.; Leveque, J.-C.; Bennett, C.; Farrokhi, F.; Piccardi, M. Machine learning applications to neuroimaging for glioma detection and classification: An artificial intelligence augmented systematic review. J. Clin. Neurosci. 2021, 89, 177–198.

14. Chow, D.; Chang, P.; Weinberg, B.D.; Bota, D.A.; Grinband, J.; Filippi, C.G. Imaging Genetic Heterogeneity in Glioblastoma and Other Glial Tumors: Review of Current Methods and Future Directions. Am. J. Roentgenol. 2018, 210, 30–38.

15. Pesapane, F.; Codari, M.; Sardanelli, F. Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. Eur. Radiol. Exp. 2018, 2, 35.

16. Pemberton, H.G.; Zaki, L.A.M.; Goodkin, O.; Das, R.K.; Steketee, R.M.E.; Barkhof, F.; Vernooij, M.W. Technical and clinical validation of commercial automated volumetric MRI tools for dementia diagnosis-a systematic review. Neuroradiology 2021, 63, 1773–1789.

17. Richens, J.G.; Lee, C.M.; Johri, S. Improving the accuracy of medical diagnosis with causal machine learning. Nat. Commun. 2020, 11, 3923.

18. Rubin, D.L. Artificial Intelligence in Imaging: The Radiologist's Role. J. Am. Coll. Radiol. 2019, 16, 1309–1317.

19. Brereton, R.G.; Lloyd, G.R. Support Vector Machines for classification and regression. Analyst 2010, 135, 230–267.

20. Sohn, C.K.; Bisdas, S. Diagnostic Accuracy of Machine Learning-Based Radiomics in Grading Gliomas: Systematic Review and Meta-Analysis. Contrast Media Mol. Imaging 2020, 2020, 2127062.

21. Gordillo, N.; Montseny, E.; Sobrevilla, P. State of the art survey on MRI brain tumor segmentation. Magn. Reson. Imaging 2013, 31, 1426–1438.

22. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444.

23. Chartrand, G.; Cheng, P.M.; Vorontsov, E.; Drozdzal, M.; Turcotte, S.; Pal, C.J.; Kadoury, S.; Tang, A. Deep Learning: A Primer for Radiologists. RadioGraphics 2017, 37, 2113–2131.

24. Ge, C.; Gu, I.Y.-H.; Jakola, A.S.; Yang, J. Deep Learning and Multi-Sensor Fusion for Glioma Classification Using Multistream 2D Convolutional Networks. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 17–22 July 2018; pp. 5894–5897.

25. Kabir-Anaraki, A.; Ayati, M.; Kazemi, F. Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. Biocybern. Biomed. Eng. 2019, 39, 63–74.

26. Ge, C.; Gu, I.Y.-H.; Jakola, A.S.; Yang, J. Deep semi-supervised learning for brain tumor classification. BMC Med. Imaging 2020, 20, 1–11.

27. Sharif, M.I.; Li, J.P.; Khan, M.A.; Saleem, M.A. Active deep neural network features selection for segmentation and recognition of brain tumors using MRI images. Pattern Recognit. Lett. 2020, 129, 181–189.

28. Hayashi, Y. Toward the transparency of deep learning in radiological imaging: Beyond quantitative to qualitative artificial intelligence. J. Med. Artif. Intell. 2019, 2, 19.

29. Topol, E.J. High-performance medicine: The convergence of human and artificial intelligence. Nat. Med. 2019, 25, 44–56.

30. Menze, B.H.; Jakab, A.; Bauer, S.; Kalpathy-Cramer, J.; Farahani, K.; Kirby, J.; Burren, Y.; Porz, N.; Slotboom, J.; Wiest, R.; et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Trans. Med. Imaging 2015, 34, 1993–2024.

31. Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Prior, F. The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. J. Digit. Imaging 2013, 26, 1045–1057.

32. Madhavan, S.; Zenklusen, J.-C.; Kotliarov, Y.; Sahni, H.; Fine, H.A.; Buetow, K. Rembrandt: Helping Personalized Medicine Become a Reality through Integrative Translational Research.

Mol. Cancer Res. 2009, 7, 157–167.

33. Rajan, P.V.; Karnuta, J.M.; Haeberle, H.S.; Spitzer, A.I.; Schaffer, J.L.; Ramkumar, P.N. Response to letter to the editor on "Significance of external validation in clinical machine learning: Let loose too early?". Spine J. 2020, 20, 1161–1162.

34. Wesseling, P.; Capper, D. WHO 2016 Classification of gliomas. Neuropathol. Appl. Neurobiol. 2018, 44, 139–150.

35. Brat, D.J.; Aldape, K.; Colman, H.; Figrarella-Branger, D.; Fuller, G.N.; Giannini, C.; Weller, M. cIMPACT-NOW update 5: Recommended grading criteria and terminologies for IDH-mutant astrocytomas. Acta Neuropathol. 2020, 139, 603–608.

36. Collins, G.S.; Reitsma, J.B.; Altman, D.G.; Moons, K.G. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Br. J. Surg. 2015, 102, 148–158.

37. Lambin, P.; Leijenaar, R.T.H.; Deist, T.M.; Peerlings, J.; de Jong, E.E.C.; van Timmeren, J.; Sanduleanu, S.; Larue, R.T.H.M.; Even, A.J.G.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. Nat. Rev. Clin. Oncol. 2017, 14, 749–762.

38. Park, J.E.; Kim, H.S.; Kim, D.; Park, S.Y.; Kim, J.Y.; Cho, S.J.; Kim, J.H. A systematic review reporting quality of radiomics research in neuro-oncology: Toward clinical utility and quality improvement using high-dimensional imaging features. BMC Cancer 2020, 20, 29.

39. Mongan, J.; Moy, L.; Kahn, C.E. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. Radiol. Artif. Intell. 2020, 2, e200029.

40. Whiting, P.F.; Rutjes, A.W.S.; Westwood, M.E.; Mallett, S.; Deeks, J.J.; Reitsma, J.B.; Leeflang, M.M.; Sterne, J.A.; Bossuyt, P.M.; QUADAS-2 Group. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann. Intern. Med. 2011, 155, 529–536.

41. Wolff, R.F.; Moons, K.G.; Riley, R.D.; Whiting, P.F.; Westwood, M.; Collins, G.S. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann. Intern. Med. 2019, 170, 51–58.

42. Andaur Navarro, C.L.; Damen, J.A.A.; Takada, T.; Nijman, S.W.J.; Dhiman, P.; Ma, J.; Collins, G.S.; Bajpai, R.; Riley, R.D.; Moons, K.G.M.; et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: Systematic review. BMJ 2021, 20, n2281.

43. Yao, A.D.; Cheng, D.L.; Pan, I.; Kitamura, F. Deep Learning in Neuroradiology: A Systematic Review of Current Algorithms and Approaches for the New Wave of Imaging Technology. Radiol. Artif. Intell. 2020, 2, e190026.

44. Olivier, M.; Asmis, R.; Hawkins, G.A.; Howard, T.D.; Cox, L.A. The Need for Multi-Omics Biomarker Signatures in Precision Medicine. Int. J. Mol. Sci. 2019, 20, 4781.

45. Subramanian, I.; Verma, S.; Kumar, S.; Jere, A.; Anamika, K. Multi-omics Data Integration, Interpretation, and Its Application. Bioinform. Biol. Insights 2020, 14, 1177932219899051.

Retrieved from https://encyclopedia.pub/entry/history/show/57041