Computer-Aided Diagnosis and Prediction in Brain Disorders

Vikram Venkatraghavan, Sebastian R. van der Voort, Daniel Bos, Marion Smits, Frederik Barkhof, Wiro J. Niessen, Stefan Klein, and Esther E. Bron

Abstract

Computer-aided methods have shown added value for diagnosing and predicting brain disorders and can thus support decision making in clinical care and treatment planning. This chapter will provide insight into the type of methods, their working, their input data –such as cognitive tests, imaging, and genetic data– and the types of output they provide. We will focus on specific use cases for diagnosis, i.e., estimating the current "condition" of the patient, such as early detection and diagnosis of dementia, differential diagnosis of brain tumors, and decision making in stroke. Regarding prediction, i.e., estimation of the future "condition" of the patient, we will zoom in on use cases such as predicting the disease course in multiple sclerosis and predicting patient outcomes after treatment in brain cancer. Furthermore, based on these use cases, we will assess the current state-of-the-art methodology and highlight current efforts on benchmarking of these methods and the importance of open science therein. Finally, we assess the current clinical impact of computer-aided methods and discuss the required next steps to increase clinical impact.

Key words Dementia, Stroke, Glioma, Cognitive impairment

1 Introduction

Computer-aided methods have major potential value for diagnosing and predicting outcomes in brain disorders such as dementia, brain cancer, and stroke. Diagnosis aims to determine the current "condition" of the patient. Prediction, or prognosis, on the other hand, aims to forecast the future "condition" of the patient. In this way, the patient's current and future condition can be estimated in a more detailed and accurate way, which opens up possibilities for better patient care and personalized medicine, with interventions tailored to the individual patient. Moreover, diagnosis and


Authors Stefan Klein and Esther E. Bron have contributed equally to this chapter

prediction are crucial not only for decision making in clinical care and treatment planning but also for managing the expectations of patients and their caregivers. This is particularly important in brain disorders as they may strongly affect life expectancy and quality of life, as symptoms of the disorder and side effects of the treatment can have a major impact on the patient's cognitive skills, daily functioning, social interaction, and general well-being. In clinical practice, diagnosis and prediction are typically performed using multiple sources of information, such as symptomatology, medical history, cognitive tests, brain imaging, electroencephalography (EEG), magnetoencephalography (MEG), blood tests, cerebrospinal fluid (CSF) biomarkers, histopathological or molecular findings, and lifestyle and genetic risk factors. These various pieces of information are integrated by the treating clinician, often in consensus with other experts at a multidisciplinary team meeting, in order to reach a final diagnosis and/or treatment plan. The aim of computer-aided methods for diagnosis and prediction is to support this process, in order to achieve more accurate, objective, and efficient decision making.

In the literature, numerous examples of computer-aided methods for diagnosis and prediction in brain disorders can be found. Most of the state-of-the-art methods use some form of machine learning to construct a model that maps (often high-dimensional) input data to the output variable of interest. There exists a large variation in machine learning technology, types of input data, and output variables. Chapters 1–6 introduced the main machine learning technologies used for computer-aided diagnosis and prediction. These include, on the one hand, classical methods such as linear models, support vector machines, and random forests, and on the other hand, deep learning methods such as convolutional neural networks and recurrent neural networks. These methods can be implemented either as classification models (estimating discrete labels) or as regression models (estimating continuous quantities), possibly specialized for survival (or "time-to-event") analysis. In addition, Chapter 17 highlights the category of disease progression modeling techniques, which could be considered as a specialized form of machine learning incorporating models of the disease evolution over time. Chapters 7–12 described the main types of input data used in machine learning for brain disorders: clinical evaluations, neuroimaging, EEG/MEG, genetics and omics data, electronic health records, and smartphone and sensor data. The current chapter focuses on the choice of the output variable, i.e., the diagnosis or prediction of interest (Fig. 1).

To illustrate the various ways in which machine learning could aid diagnosis and prediction, we focus on representative use cases
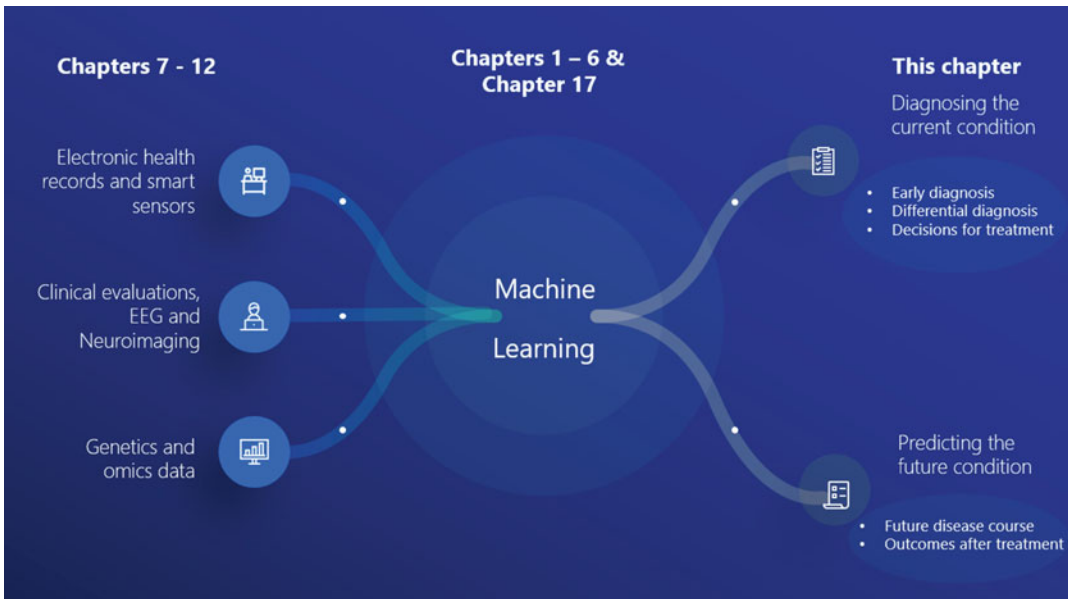
Fig. 1 Overview of the topics covered in this chapter, in the context of the other chapters in this book

organized according to the type of output. Subheading 1.1 presents diagnostic use cases, including early diagnosis, differential diagnosis, and decision making for treatment. Subheading 1.2 presents prediction use cases, including estimation of the natural disease course and prediction of patient outcomes after treatment. While the diagnostic use cases are the core of current clinical practice which could be aided by machine learning, the prediction use cases represent a potential future application. Currently, prediction is not so often made as clinicians are not yet able to make a reliable prediction in most cases. After these introductory sections, Subheading 2 provides a more comprehensive survey of the state-of-the-art methodology, and Subheading 3 analyzes the clinical impact of such methodology and suggests a roadmap for further clinical translation. Finally, Subheading 5 concludes this chapter.

1.1 Diagnosis

Diagnosis aims to determine the current "condition" of the patient to inform patient care and treatment decisions. Here, we introduce three categories of diagnostic tasks that occur in clinical practice and describe why and how computer-aided models have or could have added value.

Box 1: Diagnosis
Categories of diagnostic tasks that occur in clinical practice in which computer-aided models have or could have added value, with brain disorders for which this is relevant as examples:

- **Early diagnosis** Dementia and MS
- **Differential diagnosis** Dementia and brain cancer
- **Decision making for treatment** Stroke

Early diagnosis is highly challenging in neurodegenerative diseases such as dementia and multiple sclerosis (MS). Dementia is a clinical syndrome which can be caused by several underlying diseases, Alzheimer's disease (AD) being the most prevalent, and is estimated to affect 50 million people worldwide [4]. The mean age at dementia diagnosis is approximately 83 years [106]. MS is estimated to affect about two million people worldwide, and it primarily affects younger adults with the mean age of onset for incident MS being approximately 30 years [74]. Both for dementia and MS, establishing the diagnosis usually takes a substantial period of time after the first clinical symptoms arise [58, 139]. Early detection and accurate diagnosis is crucial for timely decision making regarding care and management of dementia symptoms, and as such can reduce healthcare costs and improve quality of life as it gives patients access to supportive therapies that help to delay institutionalization [107]. Early diagnosis of MS is important, because patients who begin treatment earlier do reap more benefit than those who start late [90]. In addition, advancing the diagnosis in time is essential to support the development of new disease-modifying treatments, since late treatment is expected to be a major factor in failure of clinical trials [88]. The clinical diagnosis of dementia is currently based on objective assessment of cognitive impairment, assessment of biomarkers [29], and evaluation of its interference with daily living [2, 42, 87, 112]. The clinical diagnosis of MS is based on frequency of relapsing inflammatory attacks, associated symptoms, and distribution of lesions on MRI [132]. For a subset of MS patients with demyelinating lesions highly suggestive of MS, termed as radiologically isolated syndrome (RIS), a separate diagnostic criteria was formed by Okuda et al. [98] to improve the diagnostic accuracy. However, objective assessment of biomarkers of the underlying processes can advance diagnosis, since symptoms are known to arise relatively late in the disease process. This holds, for example, for cognitive impairment due to dementia and physical disability or cognitive impairment due to MS [25, 40, 52]. By combining neuroimaging and other biomarkers

with machine learning based on large datasets, computer-aided diagnosis algorithms aim to facilitate medical decision support by providing a potentially more objective diagnosis than that obtained by conventional clinical criteria [63, 113]. In addition to biomarkers, machine learning based on data from remote monitoring technology, such as wearables and smart watches, is an emerging field of research aimed at detecting cognitive, behavioral, and physical symptoms in an objective way at the earliest stage possible [95, 126].

Beyond an early diagnosis, accurate identification of the underlying disease, i.e., **differential diagnosis**, is crucial for planning care and treatment decisions. For example, in dementia, the most common underlying diseases are AD, vascular cognitive impairment (VCI), dementia with Lewy bodies (DLB), and frontotemporal lobar degeneration (FTLD). Although clinical symptomatology differs between the diseases, symptoms in the early stage may be unclear and can overlap [42, 87, 112]. The current clinical criteria for AD and FTLD, for example, which entail qualitative inspection of neuroimaging, fail to accurately differentiate the two diseases [47]. Additionally, a young patient (< 65 years old) with behavioral problems could have a differential diagnosis of dementia (i.e., behavioral phenotypes of FTLD or AD) or primary psychiatric disorder, as symptomatology overlaps substantially [68]. An accurate diagnosis of primary psychiatric disorder can be informative in such patients by suggesting that progressive decline in the condition is not necessarily expected [30]. For some specific diseases, measurements of proteins causing the underlying pathology have in the last decade shown high accuracy for diagnosis of the pathology. AD is a good example with blood-based biomarkers measuring phosphorylated-Tau (P-Tau), CSF biomarkers measuring amyloid β, P-Tau and Tau, and PET imaging measuring amyloid-β and Tau. However, while highly promising, measurement of these proteins is not yet widely performed in clinical practice as blood-based biomarkers of AD are not widely available yet, CSF biomarkers require an invasive lumbar puncture, and PET imaging is too expensive and not sufficiently widely accessible to be done in each patient. Moreover, such markers of the underlying pathology are currently unavailable for other types of dementia. As an alternative, quantitative neuroimaging and other biomarkers, especially in combination with machine learning and large datasets, have shown to be beneficial in difficult cases of differential diagnosis [14, 110].

Another disorder where differential diagnosis is crucial is brain cancer. Diagnosis of brain tumors typically starts with the analysis of MRI brain data. A first diagnostic task is to differentiate between primary and secondary lesions. Primary lesions are tumors that originated from healthy brain cells, with glioma being the most common primary brain tumor type. Secondary lesions are metastases from tumors located elsewhere in the body, which may trigger

very different care and treatment paths. Also the distinction between glioma and other less common malignant primary lesions such as lymphoma is relevant. Whereas neuroradiologists are trained to differentiate these different types of lesions, the large variation in appearance of tumors induces uncertainty in the differential diagnosis. Machine learning has been shown to be able to distinguish glioma from metastasis [20] and lymphoma [86] based on quantitative analysis of brain MRI, and may thus be used as a "second" reader supporting the radiologists. Once a diagnosis of cancer is established, a second task in differential diagnosis is the further subtyping of the lesion. While glioma is one of the deadliest forms of cancer [97], there exist large differences in survival and treatment response between patients. These differences can be attributed to the glioma's genetic and histological features, in particular the isocitrate dehydrogenase (IDH) mutation status, the 1p19q co-deletion status, MGMT promoter methylation status, and the tumor grade [28, 31, 38]. These insights have led to classification guidelines by the World Health Organization (WHO) [77]. In current clinical practice, these genetic and histological features are determined from tumor tissue after resection. However, there has been an increasing interest in complementary noninvasive alternatives that can provide the genetic and histological information before resection [10, 152]. Also here, neuroradiologists can be trained to visually distinguish the subtypes based on MRI [26, 128], but uncertainty often remains and the inherent subjectivity associated with visual inspection of subtle differences in appearance, by radiologists with varying levels of expertise, is undesirable. A large body of research has therefore focused on development of machine learning approaches to support MRI-based determination of genetic and histological features of glioma [41, 65, 122, 127].

The third diagnostic task we address is **decision making for treatment**. This is relevant when multiple therapeutic options are available, such as for patients with stroke. Multiple treatment options for stroke exist such as thrombolytic medication and endovascular clot retrieval (mechanical thrombectomy). Since depending on the situation different treatments or their combination may be optimal, and since the costs per patient are rising, there is a real and urgent need for computer-aided diagnosis techniques to aid in the streamlined care of patients and individualized treatment decisions [56]. To enable early treatment of acute stroke, early and reliable diagnosis is required, which heavily relies on imaging. The vast majority of strokes are of ischemic origin, caused by a blood clot occluding an artery resulting in oxygen deprivation of the brain tissue supplied by this artery. Typical causes are large vessel occlusion with or without thrombus dislodgement (e.g., carotid stenosis) or a cardiac cause resulting in embolies (e.g., atrial fibrillation). The less common subtype is hemorrhagic stroke, which has

substantially different etiology and is often caused by hypertension. Without early treatment of stroke, prognosis is poor. Each minute without treatment leads to loss of an estimated 1.8 million neurons [64]. Patients who enter the hospital with acute stroke symptoms often immediately undergo CT (or MR) scanning, even before detailed clinical evaluation of the patient [64]. Imaging here has three roles in decision making for treatment: (1) rule out hemorrhagic stroke, (2) establish the exact cause and the extent of ischemic stroke, and (3) determine a patient's suitability for (intra-arterial) treatment [33, 80]. Applications of machine learning for treatment decisions in stroke include identification of hemorrhage and early identification of imaging findings to determine the cause and extent of stroke and estimation of the time of onset. Time of onset is relevant since most current treatments aim for rapid reperfusion of ischemic tissue, either using intravenous thrombolytic medications or using endovascular techniques to mechanically remove the obstruction to blood flow, which should be performed within 4.5 h of stroke onset [56].

1.2 Prediction

Prediction or prognosis aims to understand the future "condition" of the patient, which can then be used for considering and planning therapeutic or lifestyle interventions proactively [22] that may slow the disease process or may reduce the risk for event recurrence. In addition, it can be used for effective patient management, for managing the expectations of patients and their caregivers [82], as well as for patient selection in clinical trials [35, 102]. We distinguish two main categories of prediction targets here: the natural disease course and patient outcomes after treatment.

> **Box 2: Prediction**
> Categories of prediction targets for which computer-aided models have or could have added value, with example brain disorders for which this is relevant as discussed in this chapter:
>
> • **Natural disease course** Dementia and MS
> • **Patient outcomes after treatment** MS, brain cancer, and stroke

Predicting the natural disease course, i.e., the future progression of the disease and its symptoms in a subject, is clinically relevant as it can aid care planning and managing the expectations of patients and caregivers about their future quality of life, physical health, and dependency [81]. Additionally, in disorders where treatment options are limited, it would improve future clinical trials

for new medication through identification of patients most likely to benefit from an effective treatment, i.e., those at early stages of disease who are likely to progress over the short-to-medium term (1–5 years) [83].

In dementia, prediction is challenging because of disease heterogeneity, i.e., differences in symptoms between patients along the disease process. For example, a patient can have either typical AD with memory problems or atypical AD with either language problems [43] or behavioral problems [99]. Moreover, patients with comparable brain atrophy may decline differently as the disease progresses, reflecting cognitive resilience due to genetic or lifestyle factors that may help to compensate for the level of atrophy [147]. Lastly, a similar symptom in two patients could be resulting from different diseases altogether. For example, a patient with mild cognitive impairment (MCI) either may have early stage dementia or may have cognitive impairment due to a different cause such as older age, injury, or a virus such as SARS-CoV-2 [44]. The latter, i.e., cognitive impairment due to non-degenerative disorders, is almost twice as prevalent as cognitive impairment due to dementia [106]. Here it is of interest to predict how the symptoms will develop over time for an individual; while patients without dementia may remain stable over time or even improve, the symptoms of patients with dementia typically worsen with time. Hence, the applications of machine learning in predicting the future course of dementia include the following: (i) predicting if a patient with cognitive impairment patient will develop dementia [138], (ii) predicting when the patient will reach a clinical dementia stage (i.e., duration of the prodromal disease phase) [83], and (iii) predicting the progression of biomarkers such as cognition and MRI measurements [61, 66].

In MS, especially in the early stages when patients experience clinical symptoms sporadically, prediction of the future disease course is highly relevant for care planning and expectation management. The early stage of MS, known as the relapsing-remitting phase, is characterized by sporadic inflammatory attacks on the neuronal protective coating called myelin. Over time, the recovery from these relapses becomes incomplete, resulting in permanent and progressive disability [144]. Because of this progressive nature and the variation between individuals, predicting the number of relapses and the time to permanent disability in a specific patient is highly important for care and treatment planning [18].

Next to prediction of the natural disease course, prediction of the future disease course after an intervention, i.e., **outcome prediction after treatment**, could be instrumental for planning of treatment and subsequent follow-up. This is of particular interest in MS where multiple treatment options are available. There are

currently 21 FDA-approved disease-modifying drugs available [27] that inhibit different aspects of pathological progression of MS mainly by immune modulation and sometimes through neuroprotection or remyelination. It is hence clinically highly relevant to choose the treatment option that an individual patient is expected to have most benefit from and to determine whether risks of second-line treatment are justified [131]. The same holds for stroke in the post-acute phase, where prediction of patient outcomes after treatment based on imaging may play a role for choosing between available treatments such as medication and rehabilitation therapy [80]. Here the focus is on the long term: reducing risk of recurrence and optimization of functioning. Computer-aided approaches can thus help in personalizing the treatment for a patient.

Predicting the outcomes after treatment is also of major interest for patients with brain tumors, and specifically in case of glioma where treatment response varies greatly across patients. Treatment usually consists of surgical resection followed by radiotherapy and/or chemotherapy. Almost invariably tumor recurrence or regrowth occurs; however, the question is when. In case of high-grade glioma (i.e., glioblastoma), tumor regrowth typically happens within a few months. In low-grade glioma, progression after treatment is often slower, and it may take years before any significant regrowth is detected; at some point, however, malignant transformation (to a high-grade glioma) may occur, leading to accelerated regrowth. As discussed in Subheading 1.1, computer-aided diagnosis methods can be used to identify the current tumor's genetic and histological profile, which already provides important prognostic information. Beyond this example of computer-aided differential diagnosis, machine learning methods can contribute in different ways by directly predicting outcomes after treatment [65, 127]. First, machine learning methods have shown promise to aid the differentiation between tumor progression and treatment-related abnormalities (pseudoprogression, radiation necrosis) [54, 65, 73, 127, 143]. Second, machine learning can be used to predict local relapse locations after radiotherapy, thus highlighting locations that should be targeted with a higher radiation dose, leading to personalized radiotherapy planning [114]. Third, a machine learning approach can predict local response to stereotactic radiosurgery of brain metastases, based on radiomics analysis of pretreatment MRI, where the outcome of interest (local tumor progression) was defined in terms of maximum axial diameter growth as measured on a follow-up scan [94]. Fourth, machine learning methods have been proposed for prediction of progression-free and overall survival, which aids care planning and managing the expectations of patients about their future [60, 108, 122, 127].

2   Method Evaluation

### 2.1   State-of-the-Art Methodology for Diagnosis and Prediction

For **early diagnosis in dementia**, a large body of research has been published on classification of subjects into AD, mild cognitive impairment (MCI), and normal aging [36, 113, 145]. Overall, classification methods show high performance for classification of AD patients and cognitively normal controls with an area under the receiver operating characteristic curve (AUC) of $85 - 98\%$. Reported performances are somewhat lower for early diagnosis in patients with MCI, i.e., prediction of imminent conversion to AD (AUC: $62-82\%$). Dementia classification is usually based on clinical diagnosis as a reference standard for training and validation [87], but biological diagnosis based on assessment of amyloid pathology with PET imaging or CSF has been increasingly used over the last years [53, 129]. Structural T1-weighted (T1w) MRI to quantify neuronal loss is the most commonly used biomarker, whereas the support vector machine (SVM) is the most commonly used classifier. For T1w, both voxel-based maps (e.g., voxel-based morphometry maps quantifying local gray matter density [62]) and region-based features [78] have been frequently used. While using only region-based volumes may limit performance, combining those with regional shape and texture has been shown to perform competitively with using voxel-wise maps [13, 15, 24]. Using multimodal imaging such as FDG-PET or DTI in addition to structural MRI may have added value over structural MRI only, but limited data is available [76, 150]. Following the trends and successes in medical image analysis and machine learning, neural network classifiers –convolutional neural networks (CNN) in particular– have increasingly been used since a few years [16, 145], but have not been shown to significantly outperform conventional classifiers. In addition, data-driven disease progression models are being developed [101], which do not rely on a priori defined labels but instead derive disease progression in a data-driven way.

Regarding **differential diagnosis in dementia**, studies focus mostly on discriminating AD from other types of dementia. Differential diagnosis based on CSF and PET biomarkers of AD pathology has shown good performance for distinguishing AD from FTLD with sensitivities of 0.83 (p-tau/amyloid-β ratio from CSF) and 0.87 (amyloid PET) [48, 111, 117]. In addition, machine learning approaches have been published based on either structural or multimodal MRI as region-wise or voxel-wise imaging features and generally SVM as a classifier, similar to those used for early diagnosis in dementia. These methods focused mostly on differential diagnosis of AD and FTLD and reported performances in the range of AUC $= 0.75 - 0.85$ [12, 14, 92, 110]. A few studies addressed differential diagnosis of AD and vascular

dementia (VaD) [151] or multiclass differential diagnosis (5+ classes including AD, FTLD, VaD, dementia with Lewy bodies, and subjective cognitive decline) [93, 133].

For **differential diagnosis in brain cancer**, numerous MRI-based machine learning approaches have been presented. These developments have partly been facilitated by the availability of several valuable public datasets; see, for example, the overviews in [89, 135]. Most literature is dedicated to glioma characterization, which is therefore discussed in more detail here. Studies vary in the choice of input MRI sequences (T1w pre- and post-contrast, FLAIR, T2w, diffusion-weighted imaging, perfusion-weighted imaging, MR spectroscopy, APT CEST), the machine learning methodology (ranging from conventional radiomics approaches with hand-crafted features derived from manual tumor segmentations to deep learning approaches that automatically segment the tumor), the classification target(s) (e.g., grade, IDH, 1p19q, and/or MGMT status), the selection of glioma subtypes on which the method is validated (e.g., only low-grade glioma, only high-grade glioma, or both), and the extent of validation performed (single train-test split, repeated cross-validation, internal versus external validation). A systematic review on the use of machine learning in neuro-oncology found four articles on glioma grading, and four articles on identifying genetic/molecular characteristics of glioma based on MRI [122]. Among those, only one study used convolutional neural networks as a machine learning tool—to predict 1p19q status in low-grade glioma [1]. A more recent systematic review identified 27 studies on glioma grading of which 6 used deep learning, and 48 studies on MRI-based estimation of genetic/molecular characteristics of which 8 used deep learning [19]. Another recent review dedicated to machine learning approaches for MRI-based glioma characterization found 12 studies on glioma grading of which 2 used deep learning, and 43 studies on molecular characterization out of which 10 used deep learning [41]. These numbers indicate a trend toward deep learning approaches as we see in the entire field, but with conventional machine learning approaches with pre-defined radiomics features still being used frequently. Regarding the performance, two recent systematic reviews performed a meta-analysis of studies on molecular characterization of glioma. Jian et al. [55] found a pooled sensitivity/specificity/AUC in the validation set of 0.85/0.83/0.90 for IDH status prediction (12 studies), and 0.70/0.72/0.75 for 1p19q status prediction (5 studies). For MGMT, sensitivities and specificities ranging from 0.70 to 0.88 were found in 3 studies reporting validation performance, not allowing a meta-analysis. Van Kempen et al. [136] reported a pooled AUC of 0.91 for IDH status prediction (7 studies), 0.75 for 1p19q status prediction (3 studies), and 0.87 for MGMT promoter status prediction (3 studies). Thus, while the studies applied somewhat different

criteria for inclusion in the meta-analysis and used different statistical analysis methods, they obtained similar performance estimates. Whereas both meta-analyses suggest promising accuracy for MRI-based MGMT promoter status prediction based on the results reported in literature, a comprehensive evaluation of deep learning approaches for MGMT promoter status prediction on the BraTS2021 dataset [6] yielded disappointing results, with AUCs ranging from 0.5 to 0.6 [120]. Also, the winning method of the BraTS2021 challenge achieved an AUC of 0.62 [8], suggesting that MGMT promoter status prediction from MRI is a very difficult task. Both systematic reviews [55, 136] also pointed out the low proportion of studies with external validation (10 out of 44 in [55] and 12 out of 60 in [136]). Figure 2, recreated based on [55], shows a number of other insightful statistics on the methodologies found in literature. Finally, both reviews also identified machine learning methods aimed at predicting other, less frequently considered molecular targets, including ATRX, TERT, EGFR, P53, and PTEN, indicating the broad range of possible future research directions in this area.
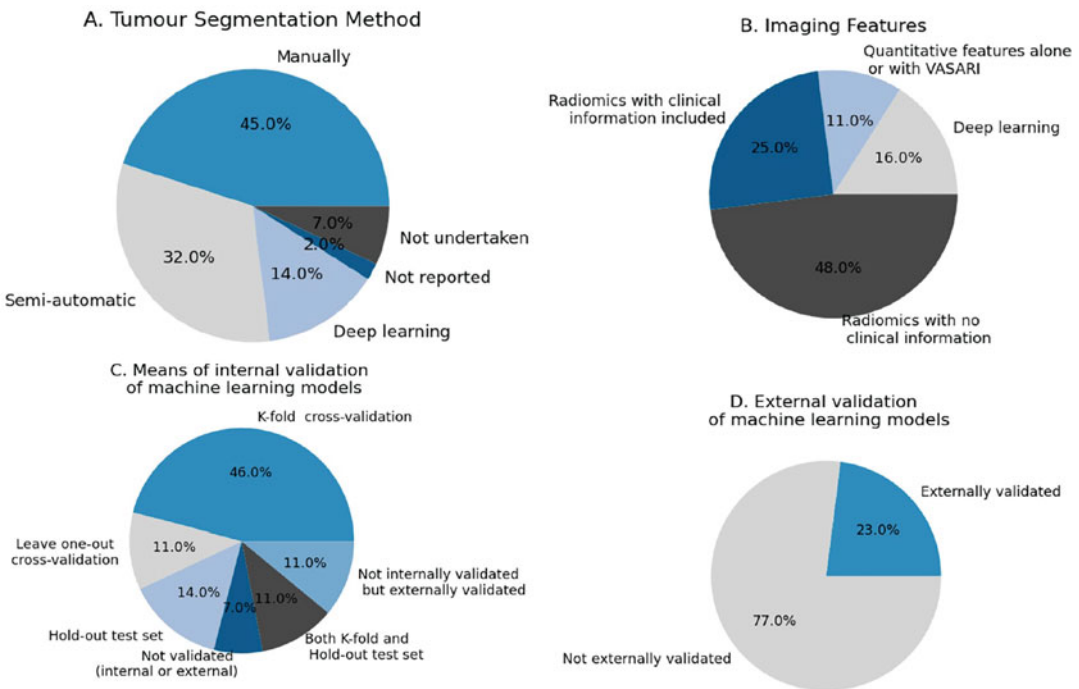

Fig. 2 Summary of tumor segmentation methods (**a**), types of imaging features (**b**), means of internal validation (**c**), and external validation (**d**) used by studies ($n = 44$) investigating machine learning models for predicting genetic subtypes of glioma. VASARI, Visually Accessible Rembrandt Imaging. Recreated from [55]. Permission to reuse was kindly granted by the publishers

Beyond glioma characterization, other differential diagnosis problems in brain cancer are differentiation between glioma and lymphoma, between glioblastoma and metastasis, between different types of meningioma, and between glioma, meningioma, and pituitary tumors [19, 65, 122, 127, 149], with promising performances reported (AUC/accuracies around 90%). Of note, a recent study pointed out an important potential source of bias (the "Clever Hans effect") in studies focused on differentiation between glioma, meningioma, and pituitary tumors, due to implicit radiologist input in the selection of the 2D slices in a commonly used benchmark dataset [142].

For decision making in stroke, different targets for machine learning based on imaging data have been identified, mostly focused at determining the cause and extent of stroke and to a lesser extent, on informing treatment decisions [56]. Regarding cause and extent of acute stroke, automatic lesion detection and identification of tissue-at-risk include the most important elements. These remain challenging as there is a lot of variation in lesion shape and location depending on time-from-symptom onset, vessel occlusion site, and collateral status [70]. Machine learning methods for segmentation and detection are increasingly successful (see Chapter 13). The step toward computer-aided diagnosis in stroke is also being taken using, for example, the CE-marked eASPECTS score [49], which is a machine learning-based assessment of the Alberta Stroke Program Early Computed Tomography Score (ASPECTS). This system for scoring acute ischemic damage to the brain has shown to be a simple, reliable, and strong predictor of functional outcome after stroke. Regarding treatment decisions, machine learning is used in several studies to determine whether a patient qualifies for a specific stroke treatment. For thrombolytic treatment, this qualification depends on time elapsed after symptom onset and treatment should be performed within 4.5 h. For this application, methods are developed that provide a binary estimation of stroke onset time (i.e., more or less than 4.5 h) based on either DWI and FLAIR [71] or perfusion-weighted imaging (CT or MR) [50]. Both approaches used a radiomics-like approach of feature extraction (e.g., intensity/gradient/texture based or using an autoencoder) followed by a machine learning classifier (support vector machine, random forest, and logistic regression). These machine learning methods had greater sensitivity than human readers using the standard procedure of DWI-FLAIR mismatch and comparable specificity. In addition, thrombolysis may cause the rare complication of symptomatic intracranial hemorrhage. Several machine learning methods have been developed to predict the risk of this complication achieving promising predictive performance, for example, using a support vector machine classifiers based on CT data (AUC $= 0.74$) [9].

For **prediction of the future course of subjects at-risk of developing dementia**, there are three frequently used approaches for defining the prediction problem at hand. First, predicting whether the patient will develop dementia. In specific diseases, measurement of proteins causing underlying pathology has shown to be very promising to identify patients in a prodromal disease state. Here, prediction is performed either using univariate analysis or using logistic regression with few variables as input. Blood-based P-Tau biomarker can predict incident AD within 4 years with an AUC of $0.78-0.83$ [103], and CSF biomarkers and PET images of amyloid β and Tau can predict clinical progression of subjects in their prodromal AD state with an AUC of $0.94-0.96$ [46]. Alternatively, in the absence of pathology-specific markers, MRI and cognitive markers of a patient together with machine learning approaches have been used to predict AD with an AUC of $0.70-0.83$ [16, 23, 75, 141]. For a systematic review of the different machine learning methods developed for the purpose of predicting AD, *see* [5]. Support vector machines (SVM) and logistic regressions are the most used algorithms in the last decade (Fig. 3). In FTD, where it is currently not possible to measure the pathological proteins in body fluids, prediction based on a combination of biomarkers that are nonspecific to the underlying pathology is promising. This is demonstrated, for example, by van der Ende et al. [134], who predicted disease onset in familial FTD based on unspecific blood-based and CSF-based biomarkers using a disease progression model and identified presymptomatic subjects that developed dementia in the near future with an AUC of 0.85.
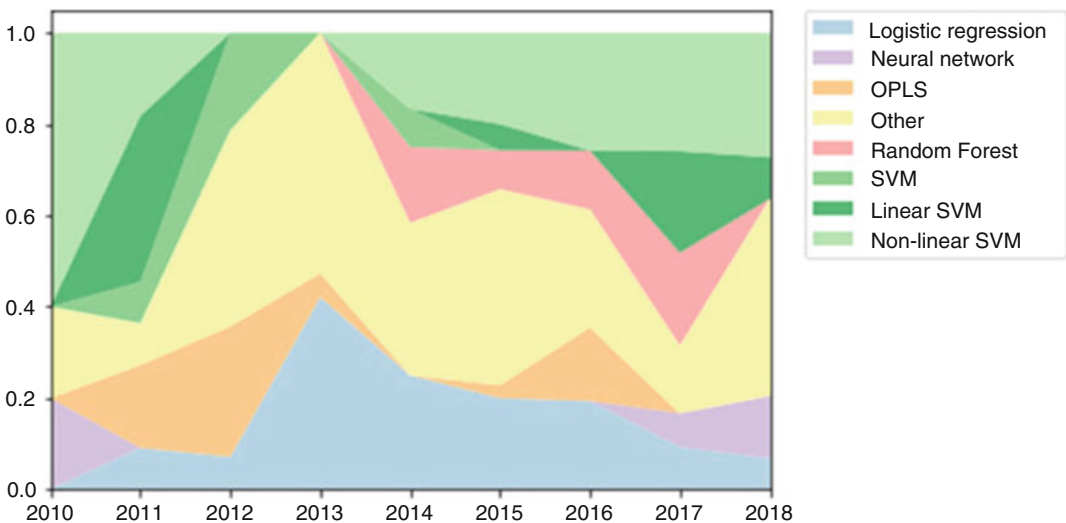
Fig. 3 Evolution with time of the use of various algorithms for predicting the progression of mild cognitive impairment. SVM with unknown kernel are simply noted as "SVM." OPLS, orthogonal partial least square; SVM, support vector machine. Reproduced from [5]. Permission to reuse was kindly granted by the publishers

Second, predicting the time for conversion to dementia. While the previous problem predicts a dichotomous output variable, here it involves predicting a continuous variable of time to dementia. Bilgel et al. [11] predicted time to AD dementia with a mean error of <1.5 years. In the TADPOLE challenge, machine learning approaches to predict time for conversion to AD dementia of 33 participating teams have been assessed quantitatively [83]. Ansart et al. [5] strongly favor predicting the exact time for conversion to dementia and argue against predicting converters within a given time interval (e.g., within 3 years), because of the precision in the predictions. While this is indeed methodologically more elegant, the implications for clinical use and perception of patients regarding prediction precision and the inherent uncertainty remain to be established.

Third, prediction of disease markers could help to obtain insight into the clinical prognosis in an individual. Important disease markers are, for example, measures of global cognition (mini-mental state examination [MMSE] or Alzheimer's disease assessment scale [ADAS] scores), or salient imaging markers (volume of the brain ventricles or longitudinal Tau protein accumulation). ADAS scores could not be reliably predicted by any participating team in the TADPOLE challenge [83], but a recent disease progression model called *AD course map* [66] could predict ADAS scores (which is scored from 0 to 150) after 3 years with a mean absolute error of 7.6 points. *AD course map* could also predict MMSE scores (which is scored from 0 to 30) after 3 years with a mean absolute error of 3.2 points. While these predictions used MRI as input, Tau PET was recently shown to be more predictive of future MMSE scores using linear mixed models [100]. However, a thorough validation of this Tau PET-based prediction is lacking. Predicting salient imaging markers such as volume of the ventricles [83], volume of the hippocampus [66], or longitudinal Tau accumulation [72] is a promising topic. Identifying the most clinically useful target to be predicted, the imaging modality that has the best cost-benefit ratio for prognosis of a patient, and the method that best predicts it are all important questions that still need answers in the future.

Most **prediction methods in MS** focus on predicting either physical disability, cognitive impairment, or treatment response in imaging data of an individual patient [57]. Physical disability as measured by expanded disability status scale (EDSS, range 0–10) has been the most commonly used predictor variable as recently used in [104, 118]. An ensemble of classifiers consisting of convolutional neural networks, random forests, and manifold learning was reported to predict EDSS with a mean square error of 3.0 [118]. Cognitive impairment has been predicted either as a global measure of cognition or as specific cognitive domains such as attention or working memory [32]. For predicting treatment

response in MS, Signori et al. [124] used meta-analysis to identify subject characteristics that have higher treatment effects. In [34], the authors used an unsupervised disease progression model to identify subtypes of progression pathways in MS and found in post hoc analysis that one of the subtypes predicted better treatment effects. Current challenges in this evolving field of predicting treatment response in MS and future directions have been summarized in [37].

For the **prediction of patient outcomes after treatment of brain cancer**, most machine learning studies have focused on MRI-based prediction of progression-free survival or overall survival, which will therefore be discussed in more detail here. A systematic review by Sarkiss et al. identified nine articles on survival prediction in glioma, and two on survival prediction for patients after stereotactic radiosurgery of brain metastases [122]. A more recent systematic review by Buchlak et al. identified 17 studies on survival prediction with performance estimates (AUC or accuracy) mostly in the range 0.7–0.8 [19]. Among those, only one study reported results of external validation, predicting overall survival of patients with low-grade glioma, and obtained an AUC of 0.71 with a model combining radiomics with non-imaging features including age, resection extent, grade, and IDH status [21]. Random (survival) forests and support vector machines were most often used methods. One study used a CNN as a pre-trained feature extractor [69]. Other recent approaches using CNNs to extract features that are subsequently combined with other factors into a final prognostic model include [45, 51, 96]. The 2017/2018 editions of the well-known BraTS challenges also included a task on overall survival prediction, with best teams obtaining accuracies around 0.6 in a three-class classification setting distinguishing short-, mid-, and long-survivors, [7]. Here, it was also pointed out that conventional machine learning methods outperformed deep learning methods, likely due to the limited size of available datasets for training.

Beyond MRI-based methods, methods using histopathology images and/or genomics data as input for the machine learning model are also considered in the literature on outcome prediction for glioma patients. In one of the pioneering studies on digital pathology images of glioma, better prognostication was obtained with deep learning when pathology images were combined with genetic markers (IDH, 1p19q) [91]. Preliminary work on so-called radiopathomics in glioma is also available, supporting the notion that combining histology and radiology features improves prognostication (overall survival prediction) in glioma patients [115, 116].

2.2 Benchmarks and Challenges

For 15 years, grand challenges have been organized in the biomedical image analysis research field. These are international benchmarks in competition form that have the goal of objectively comparing algorithms for a specific task on the same clinically

representative data using the same evaluation protocol. In such challenges, the organizers supply reference data and evaluation measures on which researchers can evaluate their algorithms. Over the past years, the number and the impact of such grand challenges have increased [79]. Also in the field of computer-aided diagnosis and prediction, such grand challenges have been organized. For example, in the dementia field, four challenges have been organized focusing on early diagnosis [3, 15, 121] and predicting the natural disease course [3, 83, 121]. In general, algorithms winning the challenges performed rigorous data pre-processing and combined a wide range of input features [17]. In the field of brain cancer, the series of BraTS challenges has had a major impact [6, 7]. These benchmarks are instrumental to gaining insight into successful approaches and their potential for use in clinical practice and clinical trials.

2.3 Open-Source Software

Open-source machine learning software such as scikit-learn[1] and MONAI[2] have been fundamental to the development of this field of research. More specifically for computer-aided diagnosis and prediction in brain diseases, dedicated platforms are available such as Clinica [119], NeuroPredict [109], and PRoNTo [123]. We also see a trend of researchers publishing their scripts and trained classifiers with their publications in order to promote reproducibility.

3 Clinical Impact

There are multiple ways in which computer-aided diagnosis and prediction models can make an impact on clinical practice. Key areas of impact are in decision making for treatment and care, replacing invasive diagnostic procedures and patient selection for clinical trials. Here we will discuss to what extent these clinical needs are addressed by current methods.

First, the most direct impact is on decision making for treatment and care. This not only affects clinical care and treatment planning in patients with, for example, dementia, stroke, MS, or brain cancer but also is important for managing the expectations of patients and their caregivers. Although high performances are achieved for some related tasks such as dementia classification, validation of those results on external datasets and clinical cohorts is still very limited as well as knowledge on the robustness of the methods. For other applications, there is still room for performance improvement, and key factors in achieving that would be the combination of multimodal input and the availability of more well-

[1] https://scikit-learn.org/.
[2] https://monai.io/.

maintained and large-scale datasets for training and evaluation. In general, there is room for improvement in how well real clinical questions are addressed by current methodology. Second, machine learning models can have an impact by replacing invasive diagnostic procedures. This is especially relevant in brain cancer, where machine learning techniques based on imaging data are developed to predict, for example, genetic mutation status or tumor grade, thereby avoiding or reducing the need for biopsies [10, 152]. As a motivating example, MRI-based prediction of MGMT methylation status could be beneficial to guide treatment decisions. This is supported by findings from a population-based study assessing survival in 131 patients with radiological diagnosis of glioblastoma who did not undergo surgery and thus lacked (histological or molecular) tissue-based verification of the diagnosis [146]. While patients without treatment had extremely poor prognosis with median survival of 3.6 months, those who received upfront temozolomide treatment did significantly better (with median survival of 6.8 months). Since the response to temozolomide is known to be highly dependent on the MGMT status, MRI-based prediction of MGMT status could give insight into which patients would benefit from treatment avoiding the need for biopsies in patients to frail for tumor biopsy. Third, patient selection for clinical trials is relevant in diseases where no to limited options for treatment exist, such as dementia, or diseases where existing treatments are suboptimal for some patients, such as MS. This can boost the power of trials by enrolling, for example, individuals who are more likely to progress based on prediction models. Several pilot studies demonstrated the added value of machine learning models to select a subgroup of participants to increase sensitivity to the treatment using phase III trial data (e.g., for Alzheimer's disease treatment using donepezil or semagacestat) [35, 102]. This will ultimately reduce the size, duration, and cost of clinical trials.

The number of published methods is not evenly distributed over tasks. While many methods have been published on, for example, the classification of Alzheimer's disease patients versus controls, much fewer publications exist on differential diagnosis in dementia. In addition, there seems to be a mismatch in some applications between published classification methods and clinical needs, e.g., the clinically relevant problem of early diagnosis does not directly translate to the frequently studied classification task of established Alzheimer's disease versus healthy controls, but would instead require separation of early disease stage Alzheimer's disease patients from those that have cognitive complaints but not dementia.

Several approved machine learning products to assist diagnosis and prediction are making their way into clinical practice, in particular in the imaging domain. Van Leeuwen et al. evaluated 100 commercially available products for AI in radiology, of which 38 are related to brain diseases [137]. These include mostly segmentation,

quantification, and normative comparison for neurodegenerative diseases and detection of lesions for stroke and oncology. Most methods generate a sample radiologist report which can be inspected and modified. In dementia, for example, 17 reporting tools that use automated brain MRI segmentation software and normative reference data for single-subject comparison are regulatory approved for use in the memory clinic [105].

One of these is Quantib ND (Quantib BV, Rotterdam, the Netherlands),[3] which is an approved commercial software that performs automatic segmentation into 20 brain regions as well as normative volumetry reference curves based on data of 5000 subjects from a population-based cohort. While Quantib ND and most other available tools use machine learning for brain segmentation, their output is not a diagnostic label produced by a machine learning algorithm. Another approved software, cDSI (Combinostics, Tampere, Finland),[4] does output diagnostic labels as confidence scores in addition to segmentation and normative volumetry based on MRI. It uses univariate machine learning to normalize individual biomarkers of different modalities based on reference values of patient and control groups, color-codes these biomarkers to improve visualization of large-data datasets, and combines confidence scores based on individual biomarkers into one score [84, 85]. While cDSI is a machine learning tool for computer-aided diagnosis and prognosis, it does not exploit the power of machine learning to detect complex patterns in high-dimensional data but rather focuses on visualization and interpretability. Diagnosis and prediction algorithms that map high-dimensional input, i.e., images and other clinical data, to an outcome measure using machine learning have not yet made their way into clinical practice.

4 Roadmap for Clinical Translation

There are numerous challenges for clinical translation of computer-aided diagnosis and prediction methods. Some key items that should be on the roadmap for translation relate to large and standardized datasets, to technical and clinical validation, to interpretability by clinicians and patients, and to practical issues related to implementation. In this section, we will discuss these requirements and related developments and initiatives.

The first requirement for translation is **large and standardized datasets**. For a few brain disorders, one or multiple large datasets (i.e., up to 2500 participants) are available to train machine learning algorithms for diagnosis and prediction tasks, facilitated

[3] quantib.com/solutions/quantib-nd.

[4] combinostics.com/cdsi.

by large multicenter initiatives such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) or the Parkinson's Progression Markers Initiative (PPMI). For validation in other cohorts and for development of algorithms in other diseases, there is only limited data available and a need for more (well-annotated) data exists. In particular, there is a need for validation data that reflect the reality of clinical routine with no to limited data harmonization and large variation in imaging protocols and data quality. Setting up such large-scale datasets is complex due to various reasons including obstacles in inter-institutional data sharing and a lack of funding for collection, curation, and labeling of data. To overcome these challenges, developments in research software and infrastructure may provide a solution by sharing easily reproducible algorithms rather than the data. Wrapping an algorithm in a container (e.g., Docker,[5] Singularity [67]) and applying the algorithms locally to the data (at one site or multiple sites in a federated approach) enables method validation on large sets of data within the confines of the local institute's firewalls. Such an approach could be also used for enabling training on larger datasets (i.e., federated learning [125]). Standardization of the data is important for eventual translation as it enables researchers to combine multiple datasets for development and validation of machine learning methods for diagnosis and prediction. Such standardization entails both data collection (e.g., diagnostic criteria, protocols for image acquisition, and clinical tests) and data organization (e.g., through open-source standards and platforms for data storage such as the Brain Imaging Data Structure (BIDS) and the Extensible Imaging Archive Toolkit (XNAT)).

Second, **technical and clinical validation** is a key focus area on the roadmap for translation. In the field of radiology, the quantitative neuroradiology initiative (QNI) framework has been developed as a model framework for translation defining the technical and clinical validation necessary to embed automated software into the clinical workflow [39]. Based on this framework, [105] reviewed the published evidence regarding commercial automated volumetric MRI tools for dementia diagnosis. For the 17 products identified, 11 companies have published some form of technical validation on their methods, but only 4 have published clinical validation in a dementia population. They concluded that there is a significant evidence gap in the literature regarding clinical validation and in-use evaluation. Whereas this review only addressed image volumetry in dementia, these findings likely extend to other brain diseases, applications, and modalities. Hence, there is a need for both retrospective and prospective studies validating algorithms in a clinical setting. In addition, performance metrics

---

[5] www.docker.com.

used in validation studies should aim to capture real clinical applicability and address different aspects of the reliability of an algorithm, including accuracy, uncertainty estimation, reproducibility, and generalizability to other data. Standards for validation and reporting are provided by guidelines such as STARD-AI [130] and TRIPOD-AI.[6]

A third key item for clinical translation is **interpretability** by end users such as clinicians and patients. As clinicians have responsibility for the decisions related to care and treatment, they should have trust in a computer-aided diagnosis or prediction system and understand its outputs to an extent that they can rely on them for decision making and explanation to a patient. Performance metrics should aim to capture real clinical applicability and be understandable to intended users [59]. High validation performance is important for building trust in methods, but not sufficient by itself, since performance may reduce in individual cases because of unaccounted inter-individual such as comorbidities or population differences such as MRI scan protocol. Therefore, apart from model accuracy, relevant questions for interpretation are, for example, Is the model suitable for the data of this patient? What features contribute to the machine learning decision for this patient? How certain is the decision for this patient and can the algorithm know when it is uncertain about an individual's decision? Such questions are important and methods should be designed and implemented in a way that facilitates answers to such questions. This could be obtained by using interpretability methods on top of "black box" machine learning models or directly by using interpretable models. For the first category, many methods have been developed based on model weight visualization, feature map visualization, backpropagation methods, or perturbation of inputs (*see* also Chapter 22). For interpretable models, an example in the field of computer-aided diagnosis and prognosis is disease progression models [140, 148]. These data-driven models are designed specifically for neurodegenerative diseases and explain their decisions based on their estimate of the natural progression of the disease in the cohort (*see* also Chapter 17).

As a final key item, we will discuss **implementation feasibility**. For machine learning models to be actually used in practice, it is essential that models and reporting are integrated into the clinical workflow and that the sending and processing of clinical data and receiving results is fully automated. Current commercial products for automatic volumetry in dementia all reported to have implemented an integration with radiology systems and the clinical workflow. While validation of the workflows is limited [105], this does support the feasibility for machine learning in clinical practice.

---

[6] osf.io/zyacb.

While these products integrate with the radiological workflow, a key challenge for the clinical translation of algorithms that use non-imaging clinical data (such as cognitive scores) as input is to also integrate with the clinical workflow of multidisciplinary diagnosis.

5   Final Summary and Conclusion

Computer-aided diagnosis and prediction of brain disorders is an important research area, with a wide variety of applications. While typically for these applications generic machine learning methods are used, domain knowledge of these brain disorders is crucial for selecting novel clinically relevant applications as well as for making domain-specific methodological improvements. Regarding diagnosis, clinical challenges are in early diagnosis of dementia and MS, differential diagnosis of dementia and brain cancer, and decision making for treatment in stroke. Regarding prediction, challenges are in the prediction of the natural disease course in dementia and MS, and the prediction of patient outcomes after treatment in stroke, brain cancer, and MS. Even though the disorders on which we focused are important avenues for impact, computer-aided diagnosis and prognosis would also be extremely useful in other disorders such as movement disorders for predicting response to treatment and side effects, epilepsy for predicting response to epilepsy surgery, and psychiatric disorders where diagnosis can be particularly difficult.

Key areas of impact are in (1) decision making for treatment and care in patients with dementia, stroke, MS, or brain cancer, (2) replacing invasive diagnostic procedures in brain cancer, and (3) patient selection for clinical trials in dementia and MS. While the first AI methods are making their way to clinical practice, diagnosis and prediction algorithms that map high-dimensional input, i.e., images and other clinical data, to an outcome measure using machine learning are not yet clinically available. To enable translation, major items on the roadmap relate to the availability of large and standardized datasets and technical and clinical validation of the developed machine learning methods. In addition, other important aspects are interpretability of the results by clinicians and patients, optimization of the diagnostic or treatment workflow in the clinic, and other practical issues related to implementation.

With this chapter, we aimed to provide a comprehensive overview, bringing together the clinical context of representative use cases of diagnosis and prediction in brain disorders and their state-of-the-art computer-aided methods. Future research should focus on bridging the identified gaps between clinical needs and the solutions brought by machine learning, to further improve decision making, treatment, and care in brain diseases.

6   Conflict of Interest

Acknowledgements

References

1. Akkus Z, Ali I, Sedlář J, Agrawal JP, Parney IF, Giannini C, Erickson BJ (2017) Predicting deletion of chromosomal arms 1p/19q in low-grade gliomas from MR images using machine intelligence. J Digit Imaging 30(4): 469–476. https://doi.org/10.1007/s102 78-017-9984-3

2. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7(3): 270–279. https://doi.org/10.1016/j. jalz.2011.03.008

3. Allen GI, Amoroso N, Anghel C, Balagurusamy V, Bare CJ et al (2016) Crowdsourced estimation of cognitive decline and resilience in Alzheimer's disease. Alzheimers Dement 12(6):645–653. https://doi. org/10.1016/j.jalz.2016.02.006

4. Alzheimer's Association Report (2020) Alzheimer's disease facts and figures. Alzheimers Dement 16(3):391–460. https://doi.org/10.1002/alz.12068

5. Ansart M, Epelbaum S, Bassignana G, Boône A, Bottani S, Cattai T, Couronné R, Faouzi J, Koval I, Louis M, Thibeau-Sutre E, Wen J, Wild A, Burgos N, Dormont D, Colliot O, Durrleman S (2021) Predicting the progression of mild cognitive impairment using machine learning: a systematic, quantitative and critical review. Med Image Anal 67: 101848. https://doi.org/10.1016/j. media.2020.101848

6. Baid U, Ghodasara S, Mohan S, Bilello M, Calabrese E et al (2021) The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification

7. Bakas S, Reyes M, Jakab A, Bauer S, Rempfler M et al (2019) Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge

8. Bakas S, Baid U, Calabrese E, Carr C, Colak E, Farahani K, Flanders AE, Kalpathy-Cramer J, Kitamura FC, Menze B, Mongan J, Prevedello L, Rudie J, Shinohara RT (2021) RSNA-MICCAI brain tumor radiogenomic classification. Link to the Kaggle challenge

9. Bentley P, Ganesalingam J, Carlton Jones AL, Mahady K, Epton S, Rinne P, Sharma P, Halse O, Mehta A, Rueckert D (2014) Prediction of stroke thrombolysis outcome using CT brain machine learning. NeuroImage Clin 4:635–640. https://doi.org/10.1016/j. nicl.2014.02.003

10. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, Allison T, Arnaout O, Abbosh C, Dunn IF, Mak RH, Tamimi RM, Tempany CM, Swanton C, Hoffmann U, Schwartz LH, Gillies RJ, Huang RY, Aerts HJWL (2019) Artificial

intelligence in cancer imaging: clinical challenges and applications. CA Cancer J Clin 69(2):127–157. https://doi.org/10.3322/caac.21552

11. Bilgel M, Jedynak BM, Initiative ADN (2019) Predicting time to dementia using a quantitative template of disease progression. Alzheimers Dement Diagn Assess Dis Monit 11(1):205–215. https://doi.org/10.1016/j.dadm.2019.01.005

12. Bouts MJRJ, Möller C, Hafkemeijer A, van Swieten JC, Dopper E, van der Flier WM, Vrenken H, Wink AM, Pijnenburg YAL, Scheltens P, Barkhof F, Schouten TM, de Vos F, Feis RA, van der Grond J, de Rooij M, Rombouts SARB (2018) Single subject classification of alzheimer's disease and behavioral variant frontotemporal dementia using anatomical, diffusion tensor, and resting-state functional magnetic resonance imaging. J Alzheimer's Dis 62(4):1827–1839. https://doi.org/10.3233/jad-170893

13. Bron EE, Steketee RM, Houston GC, Oliver RA, Achterberg HC, Loog M, van Swieten JC, Hammers A, Niessen WJ, Smits M, Klein S, for the Alzheimer's Disease Neuroimaging Initiative (2014) Diagnostic classification of arterial spin labeling and structural MRI in presenile early stage dementia. Hum Brain Mapp 35(9):4916–4931. https://doi.org/10.1002/hbm.22522

14. Bron E, Smits M, Papma J, Steketee R, Meijboom R, De Groot M, van Swieten J, Niessen W, Klein S (2016) Multiparametric computer-aided differential diagnosis of Alzheimer's disease and frontotemporal dementia using structural and advanced MRI. Eur Radiol 27(8):1–11. https://doi.org/10.1007/s00330-016-4691-x

15. Bron EE, Smits M, van der Flier WM, Vrenken H, Barkhof F et al (2015) Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. NeuroImage 111:562–579. https://doi.org/10.1016/j.neuroimage.2015.01.048

16. Bron EE, Klein S, Papma JM, Jiskoot LC, Venkatraghavan V, Linders J, Aalten P, De Deyn PP, Biessels GJ, Claassen JA, Middelkoop HA, Smits M, Niessen WJ, van Swieten JC, van der Flier WM, Ramakers IH, van der Lugt A (2021) Cross-cohort generalizability of deep and conventional machine learning for MRI-based diagnosis and prediction of Alzheimer's disease. NeuroImage Clin 31:102712. https://doi.org/10.1016/j.nicl.2021.102712

17. Bron EE, Klein S, Reinke A, Papma JM, Maier-Hein L, Alexander DC, Oxtoby NP (2021) Ten years of image analysis and machine learning competitions in dementia

18. Brown FS, Glasmacher SA, Kearns PKA, MacDougall N, Hunt D, Connick P, Chandran S (2020) Systematic review of prediction models in relapsing remitting multiple sclerosis. PLoS One 15(5):1–13. https://doi.org/10.1371/journal.pone.0233575

19. Buchlak QD, Esmaili N, Leveque JC, Bennett C, Farrokhi F, Piccardi M (2021) Machine learning applications to neuroimaging for glioma detection and classification: an artificial intelligence augmented systematic review. J Clin Neurosci Off J Neurosurg Soc Australas 89:177–198. https://doi.org/10.1016/j.jocn.2021.04.043

20. Chen C, Ou X, Wang J, Guo W, Ma X (2019) Radiomics-based machine learning in differentiation between glioblastoma and metastatic brain tumors. Front Oncol 9:806. https://doi.org/10.3389/fonc.2019.00806

21. Choi YS, Ahn SS, Chang JH, Kang SG, Kim EH, Kim SH, Jain R, Lee SK (2020) Machine learning and radiomic phenotyping of lower grade gliomas: improving survival prediction. Eur Radiol 30(7):3834–3842. https://doi.org/10.1007/s00330-020-06737-5

22. Crous-Bou M, Minguillón C, Gramunt N, Molinuevo JL (2017) Alzheimer's disease prevention: from risk factors to early intervention. Alzheimers Res Therapy 9(1):71. https://doi.org/10.1186/s13195-017-0297-z

23. Cui Y, Liu B, Luo S, Zhen X, Fan M, Liu T, Zhu W, Park M, Jiang T, Jin JS, the Alzheimer's Disease Neuroimaging Initiative (2011) Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. PLoS One 6(7):1–10. https://doi.org/10.1371/journal.pone.0021896

24. Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehéricy S, Habert MO, Chupin M, Benali H, Colliot O (2011) Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the adni database. NeuroImage 56(2):766–781. https://doi.org/10.1016/j.neuroimage.2010.06.013. Multivariate Decoding and Brain Reading

25. Dekker I, Schoonheim MM, Venkatraghavan V, Eijlers AJ, Brouwer I, Bron EE, Klein S, Wattjes MP, Wink AM, Geurts JJ, Uitdehaag BM, Oxtoby NP, Alexander DC, Vrenken H, Killestein J, Barkhof F,

Wottschel V (2021) The sequence of structural, functional and cognitive changes in multiple sclerosis. NeuroImage Clin 29: 102550. https://doi.org/10.1016/j.nicl.2020.102550

26. Delfanti RL, Piccioni DE, Handwerker J, Bahrami N, Krishnan A, Karunamuni R, Hattangadi-Gluth JA, Seibert TM, Srikant A, Jones KA, Snyder VS, Dale AM, White NS, McDonald CR, Farid N (2017) Imaging correlates for the 2016 update on WHO classification of grade II/III gliomas: Implications for IDH, 1p/19q and ATRX status. J Neurooncol 135(3):601–609. https://doi.org/10.1007/s11060-017-2613-7

27. Disease Modifying Therapies (2021) Disease modifying therapies for MS. National Multiple Sclerosis Society, New York, pp 3–21

28. Dubbink HJ, Atmodimedjo PN, Kros JM, French PJ, Sanson M, Idbaih A, Wesseling P, Enting R, Spliet W, Tijssen C, Dinjens WNM, Gorlia T, van den Bent MJ (2015) Molecular classification of anaplastic oligodendroglioma using next-generation sequencing: a report of the prospective randomized EORTC Brain Tumor Group 26951 phase III trial. Neuro-Oncology 18(3):388–400. https://doi.org/10.1093/neuonc/nov182

29. Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL et al (2014) Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. Lancet Neurol 13(6):614–629. https://doi.org/10.1016/s1474-4422(14)70090-0

30. Ducharme S, Price BH, Larvie M, Dougherty DD, Dickerson BC (2015) Clinical approach to the differential diagnosis between behavioral variant frontotemporal dementia and primary psychiatric disorders. Am J Psychiatry 172(9):827–837. https://doi.org/10.1176/appi.ajp.2015.14101248

31. Eckel-Passow JE, Lachance DH, Molinaro AM, Walsh KM, Decker PA et al (2015) Glioma groups based on 1p/19q, IDH, and TERT promoter mutations in tumors. N Engl J Med 372(26):2499–2508. https://doi.org/10.1056/NEJMoa1407279

32. Eijlers AJC, van Geest Q, Dekker I, Steenwijk MD, Meijer KA, Hulst HE, Barkhof F, Uitdehaag BMJ, Schoonheim MM, Geurts JJG (2018) Predicting cognitive decline in multiple sclerosis: a 5-year follow-up study. Brain 141(9):2605–2618. https://doi.org/10.1093/brain/awy202

33. El-Koussy M, Schroth G, Brekenfeld C, Arnold M (2014) Imaging of acute ischemic stroke. Eur Neurol 72(5–6):309–316. https://doi.org/10.1159/000362719

34. Eshaghi A, Young A, Wijeratne P, Prados F, Arnold D, Narayanan S, Guttmann C, Barkhof F, Alexander D, Thompson A, Chard D, Ciccarelli O (2021) Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. Nat Commun 12(1). https://doi.org/10.1038/s41467-021-22265-2

35. Ezzati A, Lipton RB, Alzheimer's Disease Neuroimaging Initiative (2020) Machine learning predictive models can improve efficacy of clinical trials for Alzheimer's disease. J Alzheimers Dis 74(1):55–63. https://doi.org/10.3233/jad-190822

36. Falahati F, Westman E, Simmons A (2014) Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. J Alzheimers Dis 41(3):685–708. https://doi.org/10.3233/jad-131928

37. Gasperini C, Prosperini L, Tintoré M, Sormani MP, Filippi M, Rio J, Palace J, Rocca MA, Ciccarelli O, Barkhof F, Sastre-Garriga J, Vrenken H, Frederiksen JL, Yousry TA, Enzinger C, Rovira A, Kappos L, Pozzilli C, Montalban X, De Stefano N, The MAGNIMS Study Group (2019) Unraveling treatment response in multiple sclerosis. Neurology 92(4):180–192. https://doi.org/10.1212/WNL.0000000000006810

38. Gessler F, Bernstock JD, Braczynski A, Lescher S, Baumgarten P, Harter PN, Mittelbronn M, Wu T, Seifert V, Senft C (2019) Surgery for glioblastoma in light of molecular markers: impact of resection and MGMT promoter methylation in newly diagnosed IDH-1 wild-type glioblastomas. Neurosurgery 84(1):190–197. https://doi.org/10.1093/neuros/nyy049

39. Goodkin O, Pemberton H, Vos SB, Prados F, Sudre CH, Moggridge J, Cardoso MJ, Ourselin S, Bisdas S, White M, Yousry T, Thornton J, Barkhof F (2019) The quantitative neuroradiology initiative framework: application to dementia. Br J Radiol 92(1101):20190365. https://doi.org/10.1259/bjr.20190365, pMID: 31368776

40. Gordon BA, Blazey TM, Su Y, Hari-Raj A, Dincer A et al (2018) Spatial patterns of neuroimaging biomarker change in individuals from families with autosomal dominant Alzheimer's disease: a longitudinal study. Lancet Neurol 17(3):241–250. https://doi.org/10.1016/S1474-4422(18)30028-0

41. Gore S, Chougule T, Jagtap J, Saini J, Ingalhalikar M (2021) A review of radiomics and deep predictive modeling in glioma characterization. Acad Radiol 28(11):1599–1621.

https://doi.org/10.1016/j.acra.2020.06.016

42. Gorelick PB, Scuteri A, Black SE, Decarli C, Greenberg SM et al (2011) Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the american heart association/american stroke association. Stroke 42(9):2672–2713. https://doi.org/10.1161/str.0b013e3182299496

43. Gorno-Tempini ML, Brambati SM, Ginex V, Ogar J, Dronkers NF, Marcone A, Perani D, Garibotto V, Cappa SF, Miller BL (2008) The logopenic/phonological variant of primary progressive aphasia. Neurology 71(16): 1227–1234. https://doi.org/10.1212/01.wnl.0000320506.79811.da

44. Hampshire A, Trender W, Chamberlain SR, Jolly AE, Grant JE, Patrick F, Mazibuko N, Williams SC, Barnby JM, Hellyer P, Mehta MA (2021) Cognitive deficits in people who have recovered from covid-19. EClinicalMedicine 39:101044. https://doi.org/10.1016/j.eclinm.2021.101044

45. Han W, Qin L, Bay C, Chen X, Yu KH, Miskin N, Li A, Xu X, Young G (2020) Deep transfer learning and radiomics feature prediction of survival of patients with high-grade gliomas. Am J Neuroradiol 41(1): 40–48. https://doi.org/10.3174/ajnr.A6365

46. Hansson O, Seibyl J, Stomrud E, Zetterberg H, Trojanowski JQ, Bittner T, Lifke V, Corradini V, Eichenlaub U, Batrla R, Buck K, Zink K, Rabe C, Blennow K, Shaw LM, for the Swedish Bio-FINDER study group, Initiative ADN (2018) CSF biomarkers of Alzheimer's disease concord with amyloid-β pet and predict clinical progression: a study of fully automated immunoassays in biofinder and adni cohorts. Alzheimers Dement 14(11):1470–1481. https://doi.org/10.1016/j.jalz.2018.01.010

47. Harris JM, Thompson JC, Gall C, Richardson AM, Neary D, du Plessis D, Pal P, Mann DM, Snowden JS, Jones M (2015) Do NIA-AA criteria distinguish alzheimer's disease from frontotemporal dementia? Alzheimers Dement 11(2):207–215. https://doi.org/10.1016/j.jalz.2014.04.516

48. Hellwig S, Frings L, Bormann T, Vach W, Buchert R, Meyer PT (2019) Amyloid imaging for differential diagnosis of dementia: incremental value compared to clinical diagnosis and [^{18}f]fdg pet. Eur J Nucl Med Mol Imaging 46(2):312–323. https://doi.org/10.1007/s00259-018-4111-3

49. Herweh C, Ringleb PA, Rauch G, Gerry S, Behrens L, Möhlenbruch M, Gottorf R, Richter D, Schieber S, Nagel S (2016) Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing CT scans of acute ischemic stroke patients. Int J Stroke 11(4):438–445. https://doi.org/10.1177/1747493016632244

50. Ho KC, Speier W, El-Saden S, Arnold CW (2017) Classifying acute ischemic stroke onset time using deep imaging features. AMIA Ann Symp Proc 2017:892–901

51. Huang H, Zhang W, Fang Y, Hong J, Su S, Lai X (2021) Overall survival prediction for gliomas using a novel compound approach. Front Oncol 11. https://doi.org/10.3389/fonc.2021.724191

52. Jack CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, Shaw LM, Vemuri P, Wiste HJ, Weigand SD, Lesnick TG, Pankratz VS, Donohue MC, Trojanowski JQ (2013) Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol 12(2):207–216. https://doi.org/10.1016/S1474-4422(12)70291-0

53. Jack Jr CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman DM, Jagust W, Jessen F, Karlawish J, Liu E, Molinuevo JL, Montine T, Phelps C, Rankin KP, Rowe CC, Scheltens P, Siemers E, Snyder HM, Sperling R, Contributors, Elliott C, Masliah E, Ryan L, Silverberg N (2018) NIA-AA research framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement 14(4):535–562. https://doi.org/10.1016/j.jalz.2018.02.018

54. Jang BS, Jeon SH, Kim IH, Kim IA (2018) Prediction of pseudoprogression versus progression using machine learning algorithm in glioblastoma. Sci Rep 8(1):12516. https://doi.org/10.1038/s41598-018-31007-2

55. Jian A, Jang K, Manuguerra M, Liu S, Magnussen J, Di Ieva A (2021) Machine learning for the prediction of molecular markers in glioma on magnetic resonance imaging: a systematic review and meta-analysis. Neurosurgery 89(1):31–44. https://doi.org/10.1093/neuros/nyab103

56. Kamal H, Lopez V, Sheth SA (2018) Machine learning in acute ischemic stroke neuroimaging. Front Neurol 9:945. https://doi.org/10.3389/fneur.2018.00945

57. Kanber B, Nachev P, Barkhof F, Calvi A, Cardoso J, Cortese R, Prados F, Sudre C, Tur C, Ourselin S, Ciccarelli O (2019) High-dimensional detection of imaging

response to treatment in multiple sclerosis. NPJ Digit Med 2(1). https://doi.org/10.1038/s41746-019-0127-8

58. Kaufmann M, Kuhle J, Puhan MA, Kamm CP, Chan A, Salmen A, Kesselring J, Calabrese P, Gobbi C, Pot C, Steinemann N, Rodgers S, von Wyl V, Swiss Multiple Sclerosis Registry (SMSR) (2018) Factors associated with time from first-symptoms to diagnosis and treatment initiation of multiple sclerosis in switzerland. Mult Scler J Exp Transl Clin 4(4): 2055217318814562. https://doi.org/10.1177/2055217318814562

59. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D (2019) Key challenges for delivering clinical impact with artificial intelligence. BMC Med 17(1):195. https://doi.org/10.1186/s12916-019-1426-2

60. Kickingereder P, Neuberger U, Bonekamp D, Piechotta PL, Götz M et al (2018) Radiomic subtyping improves disease stratification beyond key molecular, clinical, and standard imaging characteristics in patients with glioblastoma. Neuro-Oncology 20(6):848–857. https://doi.org/10.1093/neuonc/nox188

61. Kim J, Park Y, Park S, Jang H, Kim HJ, Na DL, Lee H, Seo SW (2021) Prediction of tau accumulation in prodromal Alzheimer's disease using an ensemble machine learning approach. Sci Rep 11(1):5706. https://doi.org/10.1038/s41598-021-85165-x

62. Klöppel S, Stonnington CM, Barnes J, Chen F, Chu C, Good CD, Mader I, Mitchell LA, Patel AC, Roberts CC, Fox NC, Jack J Clifford R, Ashburner J, Frackowiak RSJ (2008) Accuracy of dementia diagnosis–a direct comparison between radiologists and a computerized method. Brain 131(11): 2969–2974. https://doi.org/10.1093/brain/awn239

63. Klöppel S, Abdulkadir A, Jack CR, Koutsouleris N, Mourão-Miranda J, Vemuri P (2012) Diagnostic neuroimaging across diseases. NeuroImage 61(2):457–463. https://doi.org/10.1016/j.neuroimage.2011.11.002. Neuroimaging: Then, Now and the Future

64. Knight-Greenfield A, Nario JJQ, Gupta A (2019) Causes of acute stroke: a patterned approach. Radiol Clin North Am 57(6): 1093–1108. https://doi.org/10.1016/j.rcl.2019.07.007

65. Kocher M, Ruge MI, Galldiks N, Lohmann P (2020) Applications of radiomics and machine learning for radiotherapy of malignant brain tumors. Strahlenther Onkol 196(10):856–867. https://doi.org/10.1007/s00066-020-01626-8

66. Koval I, Bône A, Louis M, Lartigue T, Bottani S, Marcoux A, Samper-González J, Burgos N, Charlier B, Bertrand A, Epelbaum S, Colliot O, Allassonnière S, Durrleman S (2021) AD course map charts Alzheimer's disease progression. Sci Rep 11(1): 8020. https://doi.org/10.1038/s41598-021-87434-1

67. Kurtzer GM, Sochat V, Bauer MW (2017) Singularity: scientific containers for mobility of compute. PLoS One 12(5):1–20. https://doi.org/10.1371/journal.pone.0177459

68. Kuruppu DK, Matthews BR (2013) Young-onset dementia. Semin Neurol 33(4): 365–385. https://doi.org/10.1055/s-0033-1359320

69. Lao J, Chen Y, Li ZC, Li Q, Zhang J, Liu J, Zhai G (2017) A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme. Sci Rep 7(1):10353. https://doi.org/10.1038/s41598-017-10649-8

70. Lee EJ, Kim YH, Kim N, Kang DW (2017) Deep into the brain: artificial intelligence in stroke imaging. J Stroke 19(3):277–285. https://doi.org/10.5853/jos.2017.02054

71. Lee H, Lee EJ, Ham S, Lee HB, Lee JS, Kwon SU, Kim JS, Kim N, Kang DW (2020) Machine learning approach to identify stroke within 4.5 hours. Stroke 51(3):860–866. https://doi.org/10.1161/strokeaha.119.027611

72. Leuzy A, Smith R, Cullen NC, Strandberg O, Vogel JW, Binette AP, Borroni E, Janelidze S, Ohlsson T, Jögi J, Ossenkoppele R, Palmqvist S, Mattsson-Carlgren N, Klein G, Stomrud E, Hansson O (2021) Biomarker-based prediction of longitudinal tau positron emission tomography in Alzheimer disease. JAMA Neurol https://doi.org/10.1001/jamaneurol.2021.4654

73. Li M, Ren X, Dong G, Wang J, Jiang H, Yang C, Zhao X, Zhu Q, Cui Y, Yu K, Lin S (2021) Distinguishing pseudoprogression from true early progression in isocitrate dehydrogenase wild-type glioblastoma by interrogating clinical, radiological, and molecular features. Front Oncol 11. https://doi.org/10.3389/fonc.2021.627325

74. Liguori M, Marrosu M, Pugliatti M, Giuliani F, De Robertis F, Cocco E, Zimatore G, Livrea P, Trojano M (2000) Age at onset in multiple sclerosis. Neurol Sci J Ital Neurolog Soc Ital Soc Clin Neurophysiol 21(4 Suppl 2):S825–S829. https://doi.org/10.1007/s100720070020

75. Lin W, Tong T, Gao Q, Guo D, Du X, Yang Y, Guo G, Xiao M, Du M, Qu X, TADNI (2018)

Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment. Front Neurosci 12:777. https://doi.org/10.3389/fnins.2018.00777

76. Liu M, Cheng D, Wang K, Wang Y, Alzheimer's Disease Neuroimaging Initiative (2018) Multi-modality cascaded convolutional neural networks for alzheimer's disease diagnosis. Neuroinformatics 16(3–4):295–308. https://doi.org/10.1007/s12021-018-9370-4

77. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, Hawkins C, Ng HK, Pfister SM, Reifenberger G, Soffietti R, von Deimling A, Ellison DW (2021) The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro-Oncology 23(8):1231–1251. https://doi.org/10.1093/neuonc/noab106

78. Magnin B, Mesrob L, Kinkingnéhun S, Pélégrini-Issac M, Colliot O, Sarazin M, Dubois B, Lehéricy S, Benali H (2009) Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI. Neuroradiology 51(2):73–83. https://doi.org/10.1007/s00234-008-0463-x

79. Maier-Hein LL, Eisenmann MM, Reinke AA, Onogur SS, Stankovic MM et al (2018) Why rankings of biomedical image analysis competitions should be interpreted with care. Nat Commun 9(1). https://doi.org/10.1038/s41467-018-07619-7

80. Mair G, Wardlaw JM (2014) Imaging of acute stroke prior to treatment: current practice and evolving techniques. Br J Radiol 87(1040):20140216. https://doi.org/10.1259/bjr.20140216

81. Mank A, van Maurik IS, Bakker ED, van de Glind EM, van Der Flier W, Visser LN (2020) Identifying patient-relevant endpoints in the progression of Alzheimer's disease. Alzheimers Dement 16(S6):e040866. https://doi.org/10.1002/alz.040866

82. Mank A, van Maurik IS, Bakker ED, van de Glind EMM, Jönsson L, Kramberger MG, Novak P, Diaz A, Gove D, Scheltens P, van der Flier WM, Visser LNC (2021) Identifying relevant outcomes in the progression of Alzheimer's disease; what do patients and care partners want to know about prognosis? Alzheimers Dement Transl Res Clin Interv 7(1):e12189. https://doi.org/10.1002/trc2.12189

83. Marinescu RV, Oxtoby NP, Young AL, Bron EE, Toga AW et al (2021) The Alzheimer's disease prediction of longitudinal evolution (TADPOLE) challenge: results after 1 year follow-up. Mach Learn Biomed Imaging 1

84. Mattila J, Koikkalainen J, Virkki A, van Gils M, Lötjönen J (2012) Design and application of a generic clinical decision support system for multiscale data. IEEE Trans Biomed Eng 59(1):234–240. https://doi.org/10.1109/TBME.2011.2170986

85. Mattila J, Soininen H, Koikkalainen J, Rueckert D, Wolz R, Waldemar G, Lötjönen J (2012) Optimizing the diagnosis of early Alzheimer's disease in mild cognitive impairment subjects. J Alzheimers Dis 32(4):969–979. https://doi.org/10.3233/jad-2012-120934

86. McAvoy M, Prieto PC, Kaczmarzyk JR, Fernández IS, McNulty J, Smith T, Yu KH, Gormley WB, Arnaout O (2021) Classification of glioblastoma versus primary central nervous system lymphoma using convolutional neural networks. Sci Rep 11(1):15219. https://doi.org/10.1038/s41598-021-94733-0

87. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack Jr CR, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH (2011) The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7(3):263–269. https://doi.org/10.1016/j.jalz.2011.03.005

88. Mehta D, Jackson R, Paul G, Shi J, Sabbagh M (2017) Why do trials for Alzheimer's disease drugs keep failing? A discontinued drug perspective for 2010-2015. Expert Opin Investig Drugs 26(6):735–739. https://doi.org/10.1080/13543784.2017.1323868

89. Menze B, Isensee F, Wiest R, Wiestler B, Maier-Hein K, Reyes M, Bakas S (2021) Analyzing magnetic resonance imaging data from glioma patients using deep learning. Comput Med Imaging Graph 88:101828. https://doi.org/10.1016/j.compmedimag.2020.101828

90. Miller JR (2004) The importance of early diagnosis of multiple sclerosis. J Manag Care Pharm 10(3 Suppl B):S4–11

91. Mobadersany P, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, Velázquez Vega JE, Brat DJ, Cooper LAD (2018) Predicting cancer outcomes from histology and genomics using convolutional networks. Proc

Natl Acad Sci 115(13):E2970–E2979. https://doi.org/10.1073/pnas.1717139115

92. Möller C, Pijnenburg YAL, van der Flier WM, Versteeg A, Tijms B, de Munck JC, Hafkemeijer A, Rombouts SARB, van der Grond J, van Swieten J, Dopper E, Scheltens P, Barkhof F, Vrenken H, Wink AM (2016) Alzheimer disease and behavioral variant frontotemporal dementia: automatic classification based on cortical atrophy for single-subject diagnosis. Radiology 279(3): 838–848. https://doi.org/10.1148/radiol.2015150220

93. Morin A, Samper-Gonzalez J, Bertrand A, Ströer S, Dormont D, Mendes A, Coupé P, Ahdidan J, Lévy M, Samri D, Hampel H, Dubois B, Teichmann M, Epelbaum S, Colliot O (2020) Accuracy of MRI classification algorithms in a tertiary memory center clinical routine cohort. J Alzheimers Dis 74(4): 1157–1166. https://doi.org/10.3233/jad-190594

94. Mouraviev A, Detsky J, Sahgal A, Ruschin M, Lee YK, Karam I, Heyn C, Stanisz GJ, Martel AL (2020) Use of radiomics for the prediction of local control of brain metastases after stereotactic radiosurgery. Neuro-Oncology 22(6):797–805. https://doi.org/10.1093/neuonc/noaa007

95. Muurling M, de Boer C, Kozak R, Religa D, Koychev I, Verheij H, Nies VJM, Duyndam A, Sood M, Fröhlich H, Hannesdottir K, Erdemli G, Lucivero F, Lancaster C, Hinds C, Stravopoulos TG, Nikolopoulos S, Kompatsiaris I, Manyakov NV, Owens AP, Narayan VA, Aarsland D, Visser PJ, RADAR-AD Consortium (2021) Remote monitoring technologies in Alzheimer's disease: design of the RADAR-AD study. Alzheimers Res Therapy 13(1):89. https://doi.org/10.1186/s13195-021-00825-4

96. Nie D, Lu J, Zhang H, Adeli E, Wang J, Yu Z, Liu L, Wang Q, Wu J, Shen D (2019) Multi-channel 3d deep feature learning for survival time prediction of brain tumor patients using multi-modal neuroimages. Sci Rep 9(1): 1103. https://doi.org/10.1038/s41598-018-37387-9

97. Office for National Statistics (2019) Cancer survival in England: adult, stage at diagnosis and childhood – patients followed up to 2018. Dandy Booksellers Ltd., London

98. Okuda DT, Siva A, Kantarci O, Inglese M, Katz I et al (2014) Radiologically isolated syndrome: 5-year risk for an initial clinical event. PLoS One 9(3):1–9. https://doi.org/10.1371/journal.pone.0090509

99. Ossenkoppele R, Singleton EH, Groot C, Dijkstra AA, Eikelboom WS et al (2021) Research criteria for the behavioral variant of Alzheimer disease: a systematic review and meta-analysis. JAMA Neurol https://doi.org/10.1001/jamaneurol.2021.4417

100. Ossenkoppele R, Smith R, Mattsson-Carlgren N, Groot C, Leuzy A et al (2021) Accuracy of tau positron emission tomography as a prognostic marker in preclinical and prodromal Alzheimer disease: a head-to-head comparison against amyloid positron emission tomography and magnetic resonance imaging. JAMA Neurol 78(8):961–971. https://doi.org/10.1001/jamaneurol.2021.1858

101. Oxtoby NP, Alexander DC, EuroPOND consortium (2017) Imaging plus x: multimodal models of neurodegenerative disease. Curr Opin Neurol 30(4):371–379. https://doi.org/10.1097/wco.0000000000000460

102. Oxtoby NP, Shand C, Cash DM, Alexander DC, Barkhof F, for the Alzheimer's Disease Neuroimaging Initiative, the Alzheimer's Disease Cooperative Study (2021) Targeted screening for alzheimer's disease clinical trials using data-driven disease progression models. medRxiv. https://doi.org/10.1101/2021.01.29.21250773

103. Palmqvist S, Tideman P, Cullen N, Zetterberg H, Blennow K, Alzheimer's Disease Neuroimaging Initiative, Dage JL, Stomrud E, Janelidze S, Mattsson-Carlgren N, Hansson O (2021) Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. Nat Med 27(6):1034–1042. https://doi.org/10.1038/s41591-021-01348-z

104. Pellegrini F, Copetti M, Sormani MP, Bovis F, de Moor C, Debray TP, Kieseier BC (2020) Predicting disability progression in multiple sclerosis: insights from advanced statistical modeling. Mult Scler J 26(14):1828–1836. https://doi.org/10.1177/1352458519887343, PMID: 31686590

105. Pemberton HG, Zaki LAM, Goodkin O, Das RK, Steketee RME, Barkhof F, Vernooij MW (2021) Technical and clinical validation of commercial automated volumetric MRI tools for dementia diagnosis-a systematic review. Neuroradiology 63(11):1773–1789. https://doi.org/10.1007/s00234-021-02746-3

106. Plassman BL, Langa KM, McCammon RJ, Fisher GG, Potter GG, Burke JR, Steffens DC, Foster NL, Giordani B, Unverzagt FW, Welsh-Bohmer KA, Heeringa SG, Weir DR,

Wallace RB (2011) Incidence of dementia and cognitive impairment, not dementia in the united states. Ann Neurol 70(3):418–426. https://doi.org/10.1002/ana.22362

107. Prince M, Bryce R, Ferri C (2011) World Alzheimer Report 2011: The benefits of early diagnosis and intervention. Alzheimers Dis Int

108. Qiu X, Gao J, Yang J, Hu J, Hu W, Kong L, Lu JJ (2020) A comparison study of machine learning (random survival forest) and classic statistic (cox proportional hazards) for predicting progression in high-grade glioma after proton and carbon ion radiotherapy. Front Oncol 10:2311. https://doi.org/10.3389/fonc.2020.551420

109. Raamana PR (2017) Neuropredict: easy machine learning and standardized predictive analysis of biomarkers 1058993. https://doi.org/10.5281/zenodo.1058993

110. Raamana PR, Rosen H, Miller B, Weiner MW, Wang L, Beg MF (2014) Three-class differential diagnosis among alzheimer disease, frontotemporal dementia, and controls. Front Neurol 5:71. https://doi.org/10.3389/fneur.2014.00071

111. Rabinovici G, Rosen H, Alkalay A, Kornak J, Furst A, Agarwal N, Mormino E, O'Neil J, Janabi M, Karydas A, Growdon M, Jang J, Huang E, DeArmond S, Trojanowski J, Grinberg L, Gorno-Tempini M, Seeley W, Miller B, Jagust W (2011) Amyloid vs FDG-PET— in the differential diagnosis of AD and FTLD. Neurology 77(23): 2034–2042. https://doi.org/10.1212/WNL.0b013e31823b9c5e. ISSN 0028-3878

112. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH et al (2011) Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain 134(9):2456–2477. https://doi.org/10.1093/brain/awr179

113. Rathore S, Habes M, Iftikhar MA, Shacklett A, Davatzikos C (2017) A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. NeuroImage 155:530–548. https://doi.org/10.1016/j.neuroimage.2017.03.057

114. Rathore S, Akbari H, Doshi J, Shukla G, Rozycki M, Bilello M, Lustig RA, Davatzikos CA (2018) Radiomic signature of infiltration in peritumoral edema predicts subsequent recurrence in glioblastoma: implications for personalized radiotherapy planning. J Med Imaging 5(2):1–10. https://doi.org/10.1117/1.JMI.5.2.021219

115. Rathore S, Iftikhar MA, Gurcan MN, Mourelatos Z (2019) Radiopathomics: integration of radiographic and histologic characteristics for prognostication in glioblastoma

116. Rathore S, Nasrallah M, Mourelatos Z (2019) NIMG-76. radiopathomics: integration of radiographic and histologic characteristics for prognostication in glioblastoma. Neuro-Oncology 21(Supplement 6):vi178–vi179. https://doi.org/10.1093/neuonc/noz175.745

117. Rivero-Santana A, Ferreira D, Perestelo-Pérez L, Westman E, Wahlund LO, Sarría A, Serrano-Aguilar P (2017) Cerebrospinal fluid biomarkers for the differential diagnosis between Alzheimer's disease and frontotemporal lobar degeneration: systematic review, HSROC analysis, and confounding factors. J Alzheimers Dis 55(2):625–644. https://doi.org/10.3233/jad-160366

118. Roca P, Attye A, Colas L, Tucholka A, Rubini P et al (2020) Artificial intelligence to predict clinical disability in patients with multiple sclerosis using FLAIR MRI. Diagn Interv Imaging 101(12):795–802. https://doi.org/10.1016/j.diii.2020.05.009

119. Routier A, Burgos N, Díaz M, Bacci M, Bottani S, El-Rifai O, Fontanella S, Gori P, Guillon J, Guyot A, Hassanaly R, Jacquemont T, Lu P, Marcoux A, Moreau T, Samper-González J, Teichmann M, Thibeau-Sutre E, Vaillant G, Wen J, Wild A, Habert MO, Durrleman S, Colliot O (2021) Clinica: an open-source software platform for reproducible clinical neuroscience studies. Front Neuroinform 15:39. https://doi.org/10.3389/fninf.2021.689675

120. Saeed N, Hardan S, Abutalip K, Yaqub M (2022) Is it possible to predict mgmt promoter methylation from brain tumor MRI scans using deep learning models?

121. Sarica A, Cerasa A, Quattrone A, Calhoun V (2018) Editorial on special issue: machine learning on MCI. J Neurosci Methods 302: 1–2. https://doi.org/10.1016/j.jneumeth.2018.03.011. A machine learning neuroimaging challenge for automated diagnosis of Alzheimer's disease

122. Sarkiss CA, Germano IM (2019) Machine learning in neuro-oncology: can data analysis from 5346 patients change decision-making paradigms? World Neurosurg 124:287–294. https://doi.org/10.1016/j.wneu.2019.01.046

123. Schrouff J, Rosa M, Rondina J, Marquand A, Chu C, Ashburner J, Phillips C, Richiardi J, Mourão-Miranda J (2013) Pronto: pattern recognition for neuroimaging toolbox. Neuroinformatics 11(3):319–337. https://doi.org/10.1007/s12021-013-9178-1

124. Signori A, Schiavetti I, Gallo F, Sormani MP (2015) Subgroups of multiple sclerosis patients with larger treatment benefits: a meta-analysis of randomized trials. Eur J Neurol 22(6):960–966. https://doi.org/10.1111/ene.12690

125. Silva S, Altmann A, Gutman B, Lorenzi M (2020) Fed-biomed: a general open-source frontend framework for federated learning in healthcare. In: Albarqouni S, Bakas S, Kamnitsas K, Cardoso MJ, Landman B, Li W, Milletari F, Rieke N, Roth H, Xu D, Xu Z (eds) Domain adaptation and representation transfer, and distributed and collaborative learning. Springer International Publishing, Cham, pp 201–210

126. Simblett S, Matcham F, Curtis H, Greer B, Polhemus A, Novák J, Ferrao J, Gamble P, Hotopf M, Narayan V, Wykes T, Remote Assessment of Disease and Relapse – Central Nervous System (RADAR-CNS) Consortium (2020) Patients' measurement priorities for remote measurement technologies to aid chronic health conditions: qualitative analysis. JMIR Mhealth Uhealth 8(6):e15086. https://doi.org/10.2196/15086

127. Singh G, Manjila S, Sakla N, True A, Wardeh AH, Beig N, Vaysberg A, Matthews J, Prasanna P, Spektor V (2021) Radiomics and radiogenomics in gliomas: a contemporary update. Br J Cancer 125(5):641–657. https://doi.org/10.1038/s41416-021-01387-w

128. Smits M (2016) Imaging of oligodendroglioma. Br J Radiol 89(1060):20150857. https://doi.org/10.1259/bjr.20150857

129. Son HJ, Oh JS, Oh M, Kim SJ, Lee JH, Roh JH, Kim JS (2020) The clinical feasibility of deep learning-based classification of amyloid pet images in visually equivocal cases. Eur J Nucl Med Mol Imaging 47(2):332–341. https://doi.org/10.1007/s00259-019-04595-y

130. Sounderajah V, Ashrafian H, Golub RM, Shetty S, De Fauw J et al (2021) Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. BMJ Open 11(6). https://doi.org/10.1136/bmjopen-2020-047709

131. Stühler E, Braune S, Lionetto F, Heer Y, Jules E, Westermann C, Bergmann A, van Hövell P, Group NS, et al (2020) Framework for personalized prediction of treatment response in relapsing remitting multiple sclerosis. BMC Med Res Methodol 20. https://doi.org/10.1186/s12874-020-0906-6

132. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T et al (2018) Diagnosis of multiple sclerosis: 2017 revisions of the mcdonald criteria. Lancet Neurol 17(2):162–173. https://doi.org/10.1016/S1474-4422(17)30470-2

133. Tong T, Ledig C, Guerrero R, Schuh A, Koikkalainen J, Tolonen A, Rhodius H, Barkhof F, Tijms B, Lemstra AW, Soininen H, Remes AM, Waldemar G, Hasselbalch S, Mecocci P, Baroni M, Lötjönen J, van der Flier W, Rueckert D (2017) Five-class differential diagnostics of neurodegenerative diseases using random undersampling boosting. NeuroImage Clin 15:613–624. https://doi.org/10.1016/j.nicl.2017.06.012

134. van der Ende EL, Bron EE, Poos JM, Jiskoot LC, Panman JL et al (2021) A data-driven disease progression model of fluid biomarkers in genetic frontotemporal dementia. Brain. https://doi.org/10.1093/brain/awab382, awab382

135. van der Voort SR, Incekara F, Wijnenga MMJ, Kapsas G, Gahrmann R, Schouten JW, Tewarie RN, Lycklama GJ, Hamer PCDW, Eijgelaar RS, French PJ, Dubbink HJ, Vincent AJPE, Niessen WJ, van den Bent MJ, Smits M, Klein S (2020) Who 2016 subtyping and automated segmentation of glioma using multi-task deep learning

136. van Kempen EJ, Post M, Mannil M, Kusters B, ter Laan M, Meijer FJA, Henssen DJHA (2021) Accuracy of machine learning algorithms for the classification of molecular features of gliomas on MRI: a systematic literature review and meta-analysis. Cancers 13(11). https://doi.org/10.3390/cancers13112606

137. van Leeuwen KG, Schalekamp S, Rutten MJ, van Ginneken B, de Rooij M (2021) Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. Eur Radiol 31(6):3797–3804. https://doi.org/10.1007/s00330-021-07892-z

138. van Maurik IS, Vos SJ, Bos I, Bouwman FH, Teunissen CE et al (2019) Biomarker-based prognosis for people with mild cognitive impairment (ABIDE): a modelling study. Lancet Neurol 18(11):1034–1044. https://doi.org/10.1016/S1474-4422(19)30283-2

139. van Vliet D, de Vugt ME, Bakker C, Pijnenburg YAL, Vernooij-Dassen MJFJ, Koopmans RTCM, Verhey FRJ (2013) Time to diagnosis in young-onset dementia as compared with late-onset dementia. Psychol Med 43(2):

423–432. https://doi.org/10.1017/S00332 91712001122

140. Venkatraghavan V, Bron EE, Niessen WJ, Klein S (2019) Disease progression timeline estimation for Alzheimer's disease using discriminative event based modeling. NeuroImage 186:518–532

141. Venkatraghavan V, Dubost F, Bron EE, Niessen WJ, de Bruijne M, Klein S (2019) Event-based modeling with high-dimensional imaging biomarkers for estimating spatial progression of dementia. In: Chung ACS, Gee JC, Yushkevich PA, Bao S (eds) Information processing in medical imaging. Springer International Publishing, Cham, pp 169–180

142. Wallis D, Buvat I (2022) Clever hans effect found in a widely used brain tumour MRI dataset. Med Image Anal 77:102368. https://doi.org/10.1016/j.media.2022.102368

143. Wang K, Qiao Z, Zhao X, Li X, Wang X, Wu T, Chen Z, Fan D, Chen Q, Ai L (2020) Individualized discrimination of tumor recurrence from radiation necrosis in glioma patients using an integrated radiomics-based model. Eur J Nucl Med Mol Imaging 47(6): 1400–1411. https://doi.org/10.1007/s002 59-019-04604-0

144. Weinshenker BG (1994) Natural history of multiple sclerosis. Ann Neurol 36(S1):S6–S11. https://doi.org/10.1002/ana.410360704

145. Wen J, Thibeau-Sutre E, Diaz-Melo M, Samper-González J, Routier A, Bottani S, Dormont D, Durrleman S, Burgos N, Colliot O (2020) Convolutional neural networks for classification of alzheimer's disease: overview and reproducible evaluation. Med Image Anal 63:101694. https://doi.org/10.1016/j.media.2020.101694

146. Werlenius K, Fekete B, Blomstrand M, Carén H, Jakola AS, Rydenhag B, Smits A (2020) Patterns of care and clinical outcome in assumed glioblastoma without tissue diagnosis: a population-based study of 131 consecutive patients. PLoS One 15(2):1–14. https://doi.org/10.1371/journal.pone.0228480

147. Yao T, Sweeney E, Nagorski J, Shulman JM, Allen GI (2020) Quantifying cognitive resilience in Alzheimer's disease: the Alzheimer's disease cognitive resilience score. PLoS One 15(11):1–21. https://doi.org/10.1371/journal.pone.0241707

148. Young AL, Oxtoby NP, Daga P, Cash DM, Fox NC, Ourselin S, Schott JM, Alexander DC (2014) A data-driven model of biomarker changes in sporadic Alzheimer's disease. Brain 137(9):2564–2577

149. Zegers C, Posch J, Traverso A, Eekers D, Postma A, Backes W, Dekker A, van Elmpt W (2021) Current applications of deep-learning in neuro-oncological MRI. Ital Assoc Biomed Phys 83:161–173. https://doi.org/10.1016/j.ejmp.2021.03.003

150. Zhang D, Wang Y, Zhou L, Yuan H, Shen D (2011) Multimodal classification of Alzheimer's disease and mild cognitive impairment. NeuroImage 55(3):856–867. https://doi.org/10.1016/j.neuroimage.2011.01.008

151. Zheng Y, Guo H, Zhang L, Wu J, Li Q, Lv F (2019) Machine learning-based framework for differential diagnosis between vascular dementia and Alzheimer's disease using structural MRI features. Front Neurol 10:1097. https://doi.org/10.3389/fneur.2019.01097

152. Zhou M, Scott J, Chaudhury B, Hall L, Goldgof D, Yeom K, Iv M, Ou Y, Kalpathy-Cramer J, Napel S, Gillies R, Gevaert O, Gatenby R (2018) Radiomics in brain tumor: image assessment, quantitative feature descriptors, and machine-learning approaches. Am J Neuroradiol 39(2): 208–216. https://doi.org/10.3174/ajnr.A5391