1. Introduction

- Start by an overview on general questions related to the diagnosis of Autism
- Next, a section about brain MRI analysis in the context of Autism and methodological questions
- Finally, summary of the thesis, goals, and importance of the research.

Autism Spectrum Disorder (ASD) is a widespread condition of neurodevelopment with heterogeneous clinical presentations and unclear biological causes. While neuroimaging has provided clues into atypical neural patterns, method challenges have prevented biomarker discovery and translational insights. However, recent trends bring optimism. Open scientific data sharing has enabled unprecedented sample sizes, coinciding with advances in multivariate techniques like Machine Learning. Yet realising the full potential of these synergies requires navigating many sources of individual variability and thoughtfully applying cutting-edge analytics. This thesis aims to advance efforts to elucidate robust, generalizable MRI-based biomarkers for ASD by developing optimised pipelines integrating DL algorithms.

Dimension reduction of high-dimensional neuroimaging data holds promise for enhancing biological psychiatry. But fully realising the potential of advanced analytics requires ongoing advances in critical areas like standardisation, integration, and interpretability. This thesis contributes methodological optimizations tailored for ASD biomarkers while promoting open science and reproducibility. More broadly, this work aligns with the "spirit of the age" in computational psychiatry - embracing scale, heterogeneity, and cross-disciplinary innovation. The future of the field rests in building synergies to translate signals from diverse methodological noise into clinical insights that tangibly improve patient outcomes.

1.1. ASD description / phenotype

1.1.1. Origins and diagnosis

Autism Spectrum Disorder (ASD) is a common developmental condition that becomes apparent early in childhood [(Emerson et al., 2017), (Pierce et al., 2019)]. Autistic individuals find communicating and interacting with other people challenging and usually have restricted and rigid interests and patterns of behaviour [(*APA, Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*, 2013)]. Many Autistic individuals are very intellectually able; nevertheless, they often experience significant social [(Heasman and Gillespie, 2018), (Milton et al., 2018)], educational, and employment disadvantages [(Bird and Flint, 2019),(Heasman, 2017)]. While the prevalence of this neurodevelopmental diagnosis has been continuously increasing since it was first described in 1956 [(Zeidan et al., 2022), ("Reports on the prevalence of autism in Ireland and a review of the services for people with autism," 2018), (Christensen, 2019), (Buescher et al., 2014),] significant racial and economic disparities in autism screening and diagnosis persist [(Zeidan et al., 2022), (Guthrie et al., 2019)]. Regardless of whether there is a true increase in prevalence or simply improved recognition, the increased number of Autistic children and adults constitutes a significant challenge for parents, educators, and health professionals because ASD symptoms are heterogeneously distributed across individuals, and the biological mechanisms underlying ASD remain a mystery.

Recent research suggests that ASD has a genetic basis, with a large number genes and genetic variants implicated [(Bourgeron, 2015), (Silva et al., 2019), (Schork et al., 2019), (Miller et al., 2019), (The Brainstorm Consortium et al., 2018), (Lee et al., 2019), (Zhang et al., 2020), (Satterstrom et al., 2020), (Yoon et al., 2020), (Ruzzo et al., 2019), (Nakagawa et al., 2019), (Pagani et al., 2019), (Ecker et al., 2015)] many of which are linked with atypical brain development [(Emerson et al., 2017), (Sha et al., 2019), (Subbaraju et al., 2017),(Heinsfeld et al., 2018),(Dickie et al., 2018),(Fishman et al., 2018),(McKinnon et al., 2019),(Lake et al., 2019), (Kishida et al., 2019),(Riddle et al., 2017), (Ha et al., 2015), (Ecker et al., 2015), (Yang et al., 2016), (Haar et al., 2016), (Pereira et al., 2018), (Bedford et al., 2020), (Pagnozzi et al., 2018), (Zheng et al., 2020)]. Environmental context is also thought to be involved in the expression of genes that confer vulnerability to Autism [(Ha et al., 2015), (Ecker et al., 2015)].

Despite strong evidence of a genetic basis, the quest to understand the biological roots of ASD is made more difficult by the fact that the condition is associated with many phenotypes [(Wolfers et al., 2019)], such that two children with ASD may have very distinct sets of behavioural and social difficulties [(Milton et al., 2018), (Elibol et al., 2016)]. While these differences at the level of behaviour are very likely to reflect distinct patterns of function at the level of the brain [(Fishman et al., 2018),(McKinnon et al., 2019),(Lake et al., 2019),(Walbrin et al., 2018),(Jiang et al., 2018),(Baker et al., 2019)], no neurobiological or phenotypic subtypes of the wide spectrum of Autism have yet been identified. High rates of comorbid mental health conditions, such as anxiety and attention deficit hyperactivity disorder [(Kushki et al., 2019), (Miller et al., 2019)], further complicate diagnosis and treatment [(Kushki et al., 2019), (Silva et al., 2019), (Schork et al., 2019), (Miller et al., 2019), (Sha et al., 2019), (The Brainstorm Consortium et al., 2018), (Allsopp et al., 2019)].

Currently, Gold-Standard procedures for diagnosing ASD involve interviews and behavioural observations that must be performed by clinical specialists [(Lord et al., 1989),(Lord et al., 1994)], typically around age of 43 months(van ’t Hof et al., 2021). Although these tests permit accurate diagnoses, the specialist training involved can mean limited availability of trained clinicians, and, as a result, long waiting times for children in need of diagnostic assessment (O’Regan, 2023). While procedures for earlier diagnosis have been developed based on behavioural signs [(Emerson et al., 2017),(Zwaigenbaum et al., 2007)], this process is not widely available[(Guthrie et al., 2019),(Zuckerman et al., 2017)]. This is concerning; research shows that earlier diagnosis and intervention is associated with better outcomes in domains such as language and cognitive abilities (Clark et al., 2018; Dawson & Burner, 2011). Delays or limited access to assessment means many Autistic children may be missing out on early interventions that could be crucial in steering developmental trajectories towards optimal long-term outcomes[(O’Regan, 2023),(van ’t Hof et al., 2021), (Rogers et al., 2014), (Clark et al., 2018; Dawson & Burner, 2011)].

There are a number of factors that influence diagnosis, such as gender, comorbidities and age. These are reviewed briefly, as they are relevant to the challenge of identifying the neurobiological bases of Autism.

1.1.2. Influence of gender

While ASD is known to have a strong genetic basis, the significant gender imbalance in rates of diagnosis poses a challenge for understanding the neurobiological underpinnings.

Boys are diagnosed at a rate of 3:1 [(Loomes et al., 2017)] globally, and 10:1 in individuals without intellectual disability [(Fombonne, 2009)]. Girls meeting ASD criteria are at a higher risk of not being clinically diagnosed [(Loomes et al., 2017),(Zeidan et al., 2022)]. ASD in girls and women tends to be under or misdiagnosed because their symptoms are different to those seen in autistic males [(Kirkovski et al., 2013)] and because females with ASD appear to be better able to "camouflage" or compensate for their difficulties [(Dean et al., 2017), (Cazalis, 2017)]. Some researchers argue that current assessment practices for ASD are not appropriate for girls [(Zeidan et al., 2022), (Beggiato et al., 2017; Van Wijngaarden-Cremers et al., 2014)]. The strong gender imbalance in rates of diagnosis creates a particular challenge for understanding the impact of gender on the neurobiological bases of ASD: there are few neuroimaging datasets with a sufficient number of females to permit an analysis of the impact of gender. However, the growing number of large open-science datasets (e.g. ABIDE, the Healthy Brain Network) now provides a sufficient number of females with a diagnosis of ASD (~100) to permit a preliminary exploration of this question.

In addition to gender, high rates of comorbid conditions like ADHD and anxiety disorders further complicate the search for biomarkers, due to symptom overlap across diagnoses.

1.1.3. Comorbidities

Comorbid diagnoses of other neurological, motor, or mental health conditions are highly prevalent amongst Autistic individuals [(Gillberg, 2010),(Ecker et al., 2015), (Ghaziuddin et al., 1998), (Simonoff et al., 2008), (Gillberg et al., 2016)], although the exact rate of comorbidities varies across samples [(Ghaziuddin et al., 1998), (Simonoff et al., 2008), (Gillberg et al., 2016)].

For instance, ADHD is estimated to be present as a comorbid diagnosis in ~X% of young children with Autism [(Kushki et al., 2019), (Miller et al., 2019), (Simonoff et al., 2008)], while anxiety and mood disorders including depression are estimated to be experienced by between X and Y% of Autistic people across their lifespan [(Ghaziuddin et al., 1998), (Simonoff et al., 2008), (Gillberg et al., 2016)]. Overlapping symptomatologies [(Kushki et al., 2019)], genetics [(The Brainstorm Consortium et al., 2018)], and neuroimaging markers across conditions such as Autism, ADHD, OCD, schizophrenia, and mood disorders, raise questions about the current model of diagnostic stratification between psychiatric conditions [(Ecker et al., 2015),(Kushki et al., 2019)]. In particular, these overlaps pose a major challenge for the analysis of brain imaging data, since the goal of distinguishing ASD vs non-ASD may be misguided, since prevalent comorbidities are not taken into account. Even worse, if comorbidities are ignored in analysis, we risk identifying biomarkers that are not related to Autism at all (Ecker et al., 2015). Addressing the issue of comorbidity requires working with large databases. However, there is a risk that the available databases, even combined together, are not yet large enough or detailed enough (e.g., in terms of clinical, phenotypic, and genetic measures) to provide adequate statistical power to solve this problem.

Along with gender and comorbidities, the neurodevelopmental nature of ASD must also be considered, as clinical profiles evolve across age ranges in heterogeneous ways based on environmental influences.

1.1.4. Age and development

Autism, defined as a neurodevelopmental condition [(*APA, Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*, 2013)], is age-dependent

[(Ecker et al., 2015), (Ha et al., 2015), (Sanders, 2015), (Wolfers et al., 2019), (Van Wijngaarden-Cremers et al., 2014)].

For instance, (Elibol et al., 2016) used an unsupervised learning approach to identify disease trajectories from the electronic health records of 13,435 patients with ASD and the forum posts of 13,743 caretakers of children with ASD. They were able to identify five groups of patients sharing similar symptom profiles. Importantly, they showed that the symptomatology of each group evolved with age, and they were able to describe symptom profiles (??)it for separate age ranges from year 0 to year 15.

If we assume that clinical symptoms are mapped to specific brain alterations, then changing symptom profiles in association with age complicates our ability to detect reproducible biomarkers. A further complication is the fact that each individual grows in different living conditions and contexts. This includes environmental factors like exposure to chemicals, rural or urban environments, levels of noise, light, as well as social determinants like education, social life, standard of living, successes, failures, potential traumas, etc. Such factors can have multiple and differential influences in terms of gene expression ((Sanders, 2015), add refs), brain, and behaviour, and at different timepoints in development So, while it is clear that ASD is associated with disrupted brain development [(Emerson et al., 2017), (Ha et al., 2015), (Ecker et al., 2015)], we currently lack an understanding of the links between Autism, atypicalities in brain structure and function,, symptomatology and age. As for comorbidities, addressing the complex influences of age requires working with large databases, with deep phenotyping, which also include data on family and environmental context.

The challenges of diagnosing and subtyping ASD repeatedly underscore the need for large, deeply phenotyped datasets. Neuroimaging has traditionally lacked access to resources on this scale or methods to capitalise on such data. Small sample sizes and limited clinical measures have hindered efforts to identify reliable neural correlates.

However, the tide is turning with the emergence of open large-scale repositories like ABIDE, ABCD, UK Biobank, and Healthy Brain Network. These datasets provide unprecedented sample sizes with extensive clinical, behavioural, and environmental phenotyping beyond conventional neuroimaging resources. Paired with new multivariate methods from ML, these mega-datasets engender optimism for pattern discovery amidst immense heterogeneity.

Nevertheless, key limitations remain around standardisation and integration across resources. Assembling sufficient samples for investigating specific conditions or demographics remains non-trivial. There is also a pressing need to develop analytics that can jointly model neuroimaging, genetics, and phenotypic variables in a coherent framework.

The quest for ASD biomarkers underscores a transformative opportunity and obligation for neuroimaging. Fulfilling the long-held promise of illuminating brain-behaviour relationships may finally be possible through synergistic advances in resources and techniques. But realising this potential will rely on continued accrual of large, multi-modal datasets coupled with integrative methods to extract insights from heterogeneous signals. The future of biomarker discovery rests in embracing scale, diversity, and cross-disciplinary solutions.

1.2. ASD detection and characterisation on brain MRI data

1.2.1. Context

Consistent with the clinical picture described above, neuroimaging studies have shown that Autism Spectrum Disorders has been characterised by an atypical brain development trajectory across lifespan[(Emerson et al., 2017),(Sha et al., 2019), (Subbaraju et al., 2017), (Heinsfeld et al., 2018), (Dickie et al., 2018), (Fishman et al., 2018), (McKinnon et al., 2019), (Lake et al., 2019), (Kishida et al., 2019), (Riddle et al., 2017), (Ha et al., 2015), (Ecker et al., 2015), (Yang et al., 2016), (Haar et al., 2016), (Pereira et al., 2018), (Bedford et al., 2020),

(Pagnozzi et al., 2018), (Zheng et al., 2020)]. Studies have shown that Autistic toddlers (2-4 years) exhibit accelerated brain overgrowth and have larger brains than non-Autistic comparisons [(Ha et al., 2015),(Ecker et al., 2015)], caused by an augmentation of cortical surface area[(Zhang et al., 2020),(Nakagawa et al., 2019),(Ha et al., 2015),(Ecker et al., 2015),(Bedford et al., 2020),(Pagnozzi et al., 2018)]. This difference is not detectable later during childhood and adolescence, and is reversed during adulthood - Autistic individuals show decreased brain volumes and accelerated cortical thinning, relative to comparisons[(Ecker et al., 2015),(Yang et al., 2016),(Pereira et al., 2018),(Zheng et al., 2020)]. However, while evidence of early brain overgrowth appears compelling, the precise developmental trajectory that follows awaits further longitudinal data [Raznahan et al., https://www-sciencedirect-com.elib.tcd.ie/science/article/pii/S0006322320320230]. The initial overgrowth of grey matter in ASD is thought to affect the developmental trajectory of cortical white matter [(Ecker et al., 2015),(Pereira et al., 2018)], leading to disrupted brain wiring (connectivity). More globally, the early enlargement of the cerebral cortex [(Zhang et al., 2020),(Nakagawa et al., 2019),(Ha et al., 2015),(Ecker et al., 2015),(Bedford et al., 2020),(Pagnozzi et al., 2018)] - is thought to contribute to life-long changes and differences in brain morphology and connectivity in ASD, compared to typical brain growth[(Ha et al., 2015),(Ecker et al., 2015),(Pereira et al., 2018)]. Although the data on early brain overgrowth appear compelling, when studying brain volumes, we should be careful to take into account methodological (longitudinal vs. cross-sectional) and cohort factors that affect population norms [REF Lee et al., 2021; Raznahan et al., 2013]. In addition, research integrating genetic and environmental information is needed to better understand the influences that might provoke dysregulation of growth in ASD infants [(Raznahan et al., 2013)]. [topic sentence here]. Such studies consistently highlight Autism-related structural differences in frontotemporal and frontoparietal regions, amygdala-hippocampal complex, cerebellum, basal ganglia, anterior and posterior cingulate regions. Since these regions are - components of neural systems supporting functions such as social cognition and interaction, and emotion and behaviour regulation, dysfunction within these circuits could yield autistic traits and symptoms [(Ha et al., 2015),(Ecker et al., 2015),(Pereira et al.,

2018)]. Structural divergences in these brain regions are not specific to Autism, however. For instance, individuals with obsessive–compulsive disorder and schizophrenia share divergent morphometry in many of these regions too [(Ecker et al., 2015)]. Modelling links between brain structure and clinical phenotypes in an interpretable way would be helpful to further improve personalised medicine and patient's care.

In addition, atypical cortical folding has been observed in ASD [(Ha et al., 2015),(Ecker et al., 2015),(Yang et al., 2016),(Pereira et al., 2018)]. It is likely that the early brain overgrowth contributes to the modification of geometric features in the brain. Genetic, molecular environmental processes might transform cortical gyrification as well[(Zhang et al., 2020),(Nakagawa et al., 2019),(Ecker et al., 2015),(Bedford et al., 2020)].

The investigation of ASD differences on functional MRI data have raised a lot of replicability issues across studies [(Ha et al., 2015),(Ecker et al., 2015),(He et al., 2020)]. Small samples that under-represent ASD population, the choice of preprocessing methods, the differences in protocol of acquisitions in data-collecting sites, might bias analyses and explain the problem of reproducibility [(Ecker et al., 2015),(King et al., 2019),(He et al., 2020)]. Meaningful results include that differences in brain activation happen in regions that are structurally divergent too [(Fishman et al., 2018),(Ha et al., 2015),(Ecker et al., 2015),(Pereira et al., 2018), (Pagnozzi et al., 2018), (Holiga et al., 2019)], like, for instance, decreased activation in regions constituting the social brain network during tasks related to emotional processing or social cognition: the amygdala, the temporal–parietal junction, the insula, and the inferior frontal cortex. In addition, network analysis studies have shown disruptive connectivity compared to TDIs, poorer for long-distance regions and richer locally [(Ha et al., 2015),(Ecker et al., 2015)]. Like anatomical specificities, functional findings would not be specific to ASD only [(Sha et al., 2019)]. Thus, these results have confirmed the linkage between brain structural and functional transformations.

1.2.2. Preprocessing methods

As this brief overview of the literature shows, structural and functional MRI studies have yielded promising convergent clues as to the neurobiological bases of Autism. Yet, robust and reproducible biomarkers, either of the diagnostic category as a whole, or of subtypes (neurotypes or biotypes), have proved elusive. I have already alluded to some of the methodological challenges (e.g., small samples, cross-sectional data, changing development) that can help explain this. Yet other factors related to how MRI data is collected and analysed pose additional challenges to the search for robust and reproducible brain-based biomarkers of Autism.

In addition to the potentially confounding role of demographic factors such as age, gender, and comorbid diagnoses, brain MRI data is also subject to both systematic and unsystematic variation associated with signal and scan quality and scan sequence parameters. In particular, head motion has been identified as a particularly troublesome source of artifacts and bias [(Backhausen et al., 2016; Gilmore et al., 2019; White et al., 2018; Reuter et al., 2015),(Ha et al., 2015),(Ecker et al., 2015)].

It is crucial, therefore, to perform a complete assessment of MRI data quality prior to data processing and analysis. Yet, there is no established standard process for the selection of good MRI scans (Backhausen et al., 2016). As the number of databases and sample sizes grows [(Di Martino et al., 2014), (Di Martino et al., 2017), (Bellec et al., 2017), (Volkow et al., 2018), (Alexander et al., 2017), (Sudlow et al., 2015), (Thompson et al., 2020)], trustworthy automated methods for quality assessment are needed Performing manual quality control is time-consuming and not efficient when dealing with thousands of scans.

Furthermore, harmonisation methods exist to alleviate the confounds issue in the general case of brain MRI studies, like the COMBAT algorithm for instance [(Johnson et al., 2007)]. However, such methods may not be appropriate for situations in which we wish to understand clinical groups for which we expect to see divergent structure or function. The same concern emerges for other

preprocessing methods developed in the context of neurotypical samples. For instance, standard preprocessing pipelines applied to both structural and functional MRI data typically include transformation to a template space, such as MNI152. This template was created using anatomical scans acquired from neurotypical (white, young, Canadian) adults. The template normalisation process, through which individual structural scans are nonlinearly warped to match this template space, may therefore negatively impact the ability to detect Autism-related alterations in brain structure, introduce biases, and lead to poorer reproducibility[(Horien et al., 2022)].

On the whole, poor harmonisation of quality control methodologies and preprocessing pipelines likely contributes to inconsistent and irreproducible findings [(Heinsfeld et al., 2018),(Ecker et al., 2015),(Dadi et al., 2019)]. To address this, the field has responded with several efforts aimed at improving and harmonising methods. For example, the NiPreps project (https://www.nipreps.org/) aims to curate a set of updated and maintained tools that can be used to organise and to process neuroimaging data, such as the BIDS data ecosystem[(Gorgolewski et al., 2017), (Gorgolewski et al., 2016)]. MRIQC (Esteban et al., 2017), and fMRIPrep (Esteban et al., 2019). These efforts are vital to establishing new standards and to enhance reproducibility in the field of biological psychiatry.

1.2.3. Machine Learning approaches

Perhaps because of some of the enduring challenges to neuroimaging research, outlined above, neuroimagers have been quick to seek to take advantage of Machine Learning (ML) techniques. ML algorithms have the potential to explore a large set of features, which gives these methods a much more powerful analytical capacity than univariate approaches. Further details of ML methodologies will be provided in Chapter 2; for now, it is sufficient to note that ML algorithms can return a prediction from these features (e.g., a diagnostic label) that makes possible the development of automatic assessment tools.

Seizing on the multivariate and predictive nature of ML tools, Autism researchers have applied them to MRI data to build diagnosis prediction tools and to model brain abnormalities [(Subbaraju et al., 2017), (Lake et al., 2019), (Wolfers et al., 2019), (Riddle et al., 2017), (Ecker et al., 2015), (Pagnozzi et al., 2018), , (Dekhil et al., 2020), (Kunda et al., 2023), (Zabihi et al., 2019), (Retico et al., 2016), (Jiang et al., 2018)]. Although testing a model on a separate dataset is essential to assess the validity of a predictive algorithm, one issue with many studies, particularly the earlier ones, is that small sample sizes often led to a lack of validation on an independent dataset (Traut et al. 2021)(Pagnozzi et al., 2018).

When successfully validated on an independent dataset, ML approaches have had moderate success, with reported prediction accuracies (for Autism diagnosis) in the range of 65-75% [(Dekhil et al., 2020), (Retico et al., 2016), (Wolfers et al., 2019), (Pagnozzi et al., 2018), (Kunda et al., 2023)]. In an effort to boost accuracy through competition, (Traut et al., 2021) held an international challenge in which competing teams predicted Autism diagnosis using a large multisite dataset comprising preprocessed anatomical and functional MRI data from > 2,000 individuals. Of the 589 models submitted, the 10 best were combined and evaluated using a subset of unseen data (from one of the sites included in the main dataset), as well as data from an additional, independent acquisition site. The best ML models fusing anatomical and functional derived parameters were found to be stable and generalisable to a new independent sample, with prediction accuracies between 70% and 80%.
One observation from this effort was the fact that prediction accuracy increased with increasing sample size. Another was that while prediction accuracy for the subset of unseen data was similar to validation accuracy, accuracy for the novel site was poorer, illustrating the challenge of generalisation, particularly to new data collection sites.

Machine Learning approaches are not without their drawbacks, however. Due to the nature of the analytical procedure - the multivariate analysis of many variables - ML requires a large amount of data to converge towards an acceptable estimator. However, since the diagnostic label Autism comprises a

wider spectrum of abilities and disabilities, this is likely to be true at the level of the brain too. As outlined above, factors such as age, gender, and the presence of comorbidities further complicate this picture. Thus, there is a risk that the available databases, even combined together, are not large enough to train a ML algorithm consistently. Moreover, datasets are heterogeneous between data collecting sites, but this heterogeneity is not explicitly accounted for in the models, leading to inconsistent results between separate datasets[((Kunda et al., 2023)), (Benkarim et al., 2022), (Traut et al., 2021)]. Further, often the data on which ML algorithms are trained are usually derivatives (e.g., volumetrics, cortical thickness), computed using multiple preprocessing steps that are time-consuming and resource-demanding. Each step may add bias to the following prediction model; accumulating such biases jeopardises the accuracy of the analysis and the following interpretations (Traut et al., 2021). Recently, the persistence of these challenges has led researchers to explore whether Deep Learning approaches offer advantages over traditional Machine Learning algorithms.

1.2.4. Deep Learning approaches

Deep Learning is a domain of Machine Learning that has been increasingly and continuously updated, especially over the last three decades (LeCun et al., 2015). The idea behind Deep Learning is that a complex nonlinear function of variables relevant for a given problem (e.g., prediction of a diagnostic label) can be learned hierarchically, and that the multi-scale relationship between variables can be learned implicitly. In other words, DL allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction (LeCun et al., 2015).

One main advantage of DL algorithms compared to more traditional ML algorithms is that the preprocessing applied to input data is minimised (LeCun et al., 2015). The layers of the DL algorithm are trained to learn implicit relevant features automatically for the given problem, without the need to perform heavy feature engineering. A number of types of DL algorithm exist, including, for instance, the Multilayer Perceptron, which is a simpler type of DL architecture

(Hastie et al., 2009); Convolutional Neural Networks (Lecun et al., 1998), which are a type of architecture initially designed to analyse images; Recursive Neural Networks (Rumelhart et al., 1988), which were designed to analyse sequential data (e.g. time-series, text). Since the field is constantly evolving, new types of DL algorithms that may be revolutionary in certain fields may emerge in the future, just like the development of Transformers architecture in 2017 (Vaswani et al., 2017), which led to the development of ChatGPT. The DL algorithms that were used in this PhD project are described in detail in **Chapter 2 - Methods**.

Initial applications of DL to the prediction of ASD diagnosis from MRI have been promising [(Heinsfeld et al., 2018), (Arya et al., 2020), (Wang et al., 2020),(Hu et al., 2020), (Khosla et al., 2019), (Lu et al., 2020)(Traut et al., 2021)].
For instance, [(Wang et al., 2020),(Lu et al., 2020)] used unsupervised deep learning models to represent functional MRI information in a smaller vector of features that was used as input of a classification algorithm to predict the absence or presence of ASD. [(Arya et al., 2020),(Dekhil et al., 2020)] used both structural and functional information in order to train supervised deep learning algorithms to predict the absence or presence of ASD. We observed that accuracies obtained have been closed to 65-75%.
(Traut et al., 2021) led a competition on predicting ASD based on derived parameters of >2000 anatomical and functional brain scans. They found that DL models tended to overfit compared to ML models, meaning that the performance of the algorithm was much higher on the training set than on the testing set.

Deep Learning faces some of the same limitations as Machine Learning. Deep Learning algorithms require a sufficient amount of data to be trained and evaluated, and this amount of data is very topic-related. It is hard to determine in advance how many scans are necessary to build an accurate and robust model for a given problem. Thus, there is a risk that, despite working with thousands of scans in this project, there could still be a lack of information and that more data could be necessary. A DL algorithm is also sensitive to the input data, in particular to the bias induced by potential preprocessing steps or by defective scans. Compared to ML, DL has additional challenges due to the

nature of the architecture of the models, in particular the fact that the number of parameters to optimise is huge. The issue of overfitting can appear like in (Traut et al., 2021), and the model built is not generalisable in that case. Furthermore, training an algorithm with millions of parameters and on a large dataset is physically challenging for a computer. Sophisticated strategies like batch optimization, parallel and distributed computing can limit the number of computations to perform and save in one iteration step, enabling the training loop to iterate well without memory issues. The choice of the type of model and architecture has also a positive or negative impact on this amount of computations too.

Lastly, another important concern is the reproducibility of DL methods. The complexity of DL algorithms can lead to the generation of more complex code. It can be challenging to openly share code that is sufficiently clear and reusable for the wider community. Sharing the optimal parameters for a particular model is also important to enable other researchers to launch a model directly on a new dataset, or to serve as a baseline called a “pre-trained model” that is then tuned to be optimised for a new task.

Addressing these challenges is a central aim of this PhD project.

1.3. PhD project

1.3.1. Importance of the proposed research

Building a Deep Learning algorithm that is interpretable biologically and returns quantitative measures of ASD brain specificities, linked with visible autistic behaviours and symptoms, would be ideal to further develop accurate precision medicine methods for ASD care.

This research project could be useful for medical practitioners, educators, parents who take care of autistic people, and also for autistic people themselves. It could also be a step towards making the field of Psychiatry evolve again thanks to neuroimaging and AI technology advances.

While many autistic people are intellectually able, most of them have encountered discrimination and misunderstanding of their attitude at school, at work, and even within their families [(Heasman and Gillespie, 2018), (Milton et al., 2018),(Bird and Flint, 2019),(Heasman, 2017)]. Improving the knowledge of their functioning at the brain level can help to tackle the unfairness of the situation and improve the tolerance and empathy for these human beings. It can also help in diagnosing ASD, especially if predictive tools are built in that sense.

Better understanding the brain differences between people diagnosed with ASD and people not diagnosed with ASD can lead to better characterise ASD at a young age. Early intervention and adaptation of the environment could be initiated, and hence leading to better progression [(van ’t Hof et al., 2021), (Rogers et al., 2014), (Clark et al., 2018; Dawson & Burner, 2011)].

In addition, this research could also be of great significance for the intervention. Establishing interventions and synchronising all the parties of this intervention (psychologists, speech therapists, teachers, educators, …) is uneasy. MRI could be a vector of understanding how an intervention is going, how the brain evolves with new daily behavioural exercises for instance. Accordingly, developing methods to analyse brain MRI data in an effective and robust way could serve in the development of such tools.

At a bigger scale, the field of psychiatry is evolving thanks to the integration of many other disciplines like neuroimaging, artificial intelligence, genetics, immunology. Many overlaps in genetics, symptoms, brain patterns have been observed between various diagnoses in Psychiatry. This project can also help the characterization of diagnoses in Psychiatry to move forward. A key stake is to reach a more precise medicine while keeping the tools efficient for the practitioners. In parallel, the project also makes the emerging field of advanced AI applied to Health and Medicine progress.

1.3.2. Research problems and questions

Building up the contextual foundations through the various sections of this introduction, we have raised many questions and issues around the topic.

What are the brain patterns that characterise Autism? Or what are the brain patterns that enable us to distinguish people with a diagnosis of Autism from people with no diagnosis of Autism? Would the patterns be structural or functional or both?
Answering these global questions will be anyhow possible only if a sufficient amount of data is available. And knowing how many patients should be in a dataset to build a relevant model is also a question. The insights from previous studies are that considering only one data-collecting site or a few dozens of patients is not sufficient to build an extensive and replicable model. Dealing with data acquired from different collecting sites is also challenging due to the presence of a multi-site effect that has been underlined in previous studies on such databases.
Moreover, the contextual information about the topic makes the topic itself questionable: is it possible to build an extensive model that detects Autism on brain MRI data while many confounds - age, gender, comorbidities - can substantially have an impact on the brain patterns observed? This issue is strengthened knowing that ASD is a large spectrum of behaviours and sensibilities that may be reflected at different areas and levels in the brain.

Regarding the technical aspects, concerns have arisen from the preprocessing pipelines and the Deep Learning models. Can we build a better preprocessing pipeline on MRI data to be able to study more accurately the brain characteristics of ASD? Can we find ways to understand, to explain or to interpret the decisions of Deep Learning models? Further, is it possible to obtain a description of brain imaging based markers of ASD by building a predictive model of the diagnosis of ASD?
The challenge posed by Deep Learning is also to have sufficient power and memory to be able to train the models, as well as to choose good strategies to optimise the models at a lower computational cost.

Another issue related to the neuroscience field is the reproducibility of research projects. In the case of this transversal PhD project, how can the code be shareable to the wider community? What are the existing current standards in neuroimaging to enhance reusability ?

The research project moved steps toward addressing these questions and challenges across the various works done.

1.3.3. Research aims and objectives

The overall objective of the PhD project was to advance research on MRI-based biomarkers of Autism by designing new analytical pipelines that include Deep Learning algorithms.

To achieve this aim, I used several large Open Science databases (described in Chapter X), and have openly shared all code, to maximise the value of this work for the community. Determining what data is available is crucial to understand what questions among all the potential ones could be studied and answered statistically.

Other goals were to:

- Improve existing pipelines for MRI data preprocessing;
- Create predictions from two modalities of MRI data: structural data and resting-state functional data;
- Tailor new Deep Learning pipelines to each MRI modality;
- Boost the acceptability of Deep Learning applications in medicine by improving the explainability and interpretability of models. This was achieved by finding and implementing methods to explain Deep Learning models and to interpret the brain patterns driving to prediction outcomes;
- Participate in and contribute to Open Science;
- Help raise the current standards of reproducibility for neuroimaging research;
- Share high quality, readable code for better reusability;

- Grow as a young researcher by developing my skills in psychiatry, Deep Learning, communication, leadership, management, and, more globally, in topics related to technology and Health.

1.3.4. Hypotheses

Main hypotheses emerged from the insights given by the context on the topic:

- Structural MRI data is precise enough to build a predictive model of ASD;
- Functional MRI data is precise enough to build a predictive model of ASD;
- Brain is a relevant variable to look at to study ASD;
- Autistic people share common characteristic patterns in the brain while being at different levels of the spectrum of Autism;
- The variability of the data is sufficient to build a robust algorithm;
- We have a sufficient amount of data.

One hypothesis that could help in the project if true would be that the variability of the brains of people with no diagnosis is greater than the variability of the brains of autistic people. Indeed, the models would need less data to converge to an extensive model. Unfortunately, in view of the lack of replicability and of specificity of brain patterns associated with ASD in the previous studies, stating this hypothesis does not seem reasonable, even if such issues could also be due to confounds such as multi-site effect, age, gender, …

1.3.5. Plan of the thesis/ Introduction of the studies

The work performed for this thesis is describe in five chapters:

- Chapter X describes my first empirical study, which aimed to build a fast, reliable quality control pipeline for brain structural MRI data using Deep Learning. The best-performing algorithm was integrated into an open BIDS-app, which was shared with the neuroimaging community.
- Chapter X aimed to build an interpretable pipeline for the prediction of detection of ASD diagnosis from structural MRI data using Deep Learning. Key innovations were (1) the model was trained on minimally

preprocessed data (no registration to template), (2) the characterization of regions that contributed to the prediction of ASD (interpretability), and (3) the examination of how age, gender, and Comorbidities influenced the characterization of such regions.

- My third empirical study is described in Chapter X3. This project aimed to build a new Deep Learning approach to prediction from resting state fMRI data; we applied it to the detection of ASD, but also to gender, age, and performed an analysis of the brain areas that contribute to prediction outcomes.
- Chapter X describes efforts aimed at promoting reproducibility and fostering better practices in neuroimaging research.
- Finally, Chapter X summarises my “extra-curricular” projects - the summer schools I attended during the PhD, and several projects undertaken for these school programmes, which helped me grow as a young researcher in psychiatry and in AI.