1. Introduction

Autism Spectrum Disorder (ASD) is a prevalent neurodevelopmental condition characterised by heterogeneous clinical presentations and unclear biological underpinnings. While neuroimaging has provided clues into atypical neural patterns, methodological challenges have hindered biomarker discovery and translational insights. Recent trends, however, engender 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 myriad 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 evolving *zeitgeist* (“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 neurodevelopmental condition characterised by social communication challenges and restrictive, repetitive behaviours [1,2]. The prevalence of ASD diagnoses has risen steadily over recent decades [3-5], posing significant challenges for families, educators, and clinicians. While many individuals with ASD are intellectually able, they face disadvantages in social, educational, and vocational outcomes [6,7].

ASD is highly heterogeneous, with a strong genetic basis but unclear biological mechanisms [8-16]. Environmental factors likely interact with genetic risks [17,18]. Phenotypic variation is high, as no behavioral or biological subtypes have been firmly identified [19-21]. Co-occurring conditions like anxiety further complicate the picture [22-24].

Gold standard ASD diagnosis relies on specialist behavioural assessments around age 3 years [25,26]. However, limited availability of experts leads to long wait times [27]. Earlier screening methods exist but are not widely implemented [28,29], despite evidence that early intervention improves outcomes [30,31].

Factors like gender, comorbidities, and age-related changes challenge efforts to elucidate ASD's neurobiological roots. For example, the marked gender imbalance in diagnosis (3-10:1 male predominance) [32] hinders study of gender influences on ASD neurobiology [33,34]. High comorbidity confounds attempts to identify ASD-specific neural correlates [22,24]. And symptom profiles and neural patterns change across development in heterogeneous ways [35-37]. However, in this thesis

project, it was assumed that Autistic people share common characteristic patterns in the brain while being at different levels of the spectrum of Autism, based on the global assumption that Brain was a relevant variable to look at to study ASD.

1.1.2. Influence of gender

The significant gender imbalance in ASD diagnosis poses challenges for understanding neurobiological factors. ASD is diagnosed at 3:1 (boys:girls) globally, and 10:1 without intellectual disability [1,2]. Girls meeting diagnostic criteria often go unidentified or misdiagnosed due to differing symptom presentation from boys and increased ability to mask difficulties [3-5]. Current assessment practices may be ill-suited for girls [6,7]. The scarcity of diagnosed females hinders neuroimaging research on gender influences. However, growing open-science datasets now provide sufficient female samples (~100 with ASD) to enable preliminary investigation [8].

High comorbidity rates further confound the search for ASD biomarkers. Co-occurring conditions like ADHD and anxiety share genetic risks, brain patterns, and symptom overlap with ASD [9,10]. Disentangling disorder-specific neural correlates becomes challenging [11]. Large datasets with deep phenotyping are needed to address this but may still lack statistical power currently [12].

From previous work, questioning arises:

- How does ASD manifest differently in males vs. females, both behaviorally and neurally?

- What factors contribute to underdiagnosis of females, and how can assessment practices be adapted accordingly?
- Do neuroimaging biomarkers differ between males and females with ASD? If so, how does this inform our understanding of gender-dependent neurophenotypes?

1.1.3. Comorbidities

High rates of comorbid conditions are common in ASD, including ADHD (~30-50%), anxiety, depression (30-70%), and other neurological or developmental disorders [1-5]. However, exact comorbidity rates vary across studies [2,3,5].

Extensive symptomatic, genetic, and neural overlap exists between ASD and conditions like ADHD, OCD, and schizophrenia [6-8]. This challenges the current model of discrete psychiatric diagnoses [7,8] and complicates identification of ASD-specific biomarkers. Analyses ignoring comorbidity risk finding non-specific correlates [8].

Large datasets with deep phenotyping are required to address comorbidities. However, available resources may still lack adequate power due to extensive heterogeneity in ASD [8].

Research questions can be drawn about ASD and comorbidities:

- To what extent do comorbid conditions like ADHD and anxiety share common genetic and neural correlates with ASD?

- How can neuroimaging analyses disentangle disorder-specific biomarkers from overlapping correlates in the presence of comorbidities?
- What sample sizes and depth of phenotypic data are required to unravel the complex relationships between ASD, comorbidities, and neurobiology?

1.1.4. Age and development

Autism is a neurodevelopmental disorder with clinical profiles changing across the lifespan [1-5]. For example, an analysis of electronic records identified distinct, age-dependent ASD symptom trajectories from ages 0-15 years [3].

If symptoms map to specific neural patterns, these shifting profiles over development likely complicate detecting robust biomarkers [4]. Each individual also grows up in unique environmental contexts influencing brain and behaviour [1,2,5,6]. Such factors include toxins, noise, social determinants, and potential traumas [5,6].

While ASD involves atypical neurodevelopment [7-12], precisely linking symptoms, brain changes, and age remains limited. Large datasets with deep phenotyping of environmental and developmental contexts are needed [2,4].

Given the evolutive aspect of ASD, several questions come:

- How do neural correlates of ASD evolve over development from early childhood to adulthood?

- What environmental variables interact with genetics to shape developmental trajectories in ASD?
- Can longitudinal neuroimaging identify critical periods of divergence or convergence in brain development?
- How can changing age-related symptom profiles be linked to underlying neural dynamics over the lifespan?

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.

By taking advantage of these large-scale database projects, I hypothesised that there was a sufficient amount of data with a large enough variability to build robust algorithms.

The quest for ASD biomarkers underscores a transformative opportunity for neuroimaging. Fulfilling the long-held promise of illuminating brain-behaviour relationships may finally be possible through synergistic advances in resources and techniques.

1.2. ASD detection and characterisation on brain MRI data

1.2.1. Context

Neuroimaging studies reveal an atypical developmental trajectory in ASD [1-10]. Autistic toddlers exhibit accelerated brain overgrowth and enlarged cortical surface area [4,5,7-10], which reverses by adulthood with decreased brain volume and accelerated cortical thinning [3,6,9]. While compelling, precise characterization of this trajectory awaits further longitudinal data [11].

The initial overgrowth may disrupt white matter development and contribute to altered morphology and connectivity lifelong [3,5,9]. However, methodological and cohort factors influencing population norms need consideration when interpreting volumetrics [12,13]. Integrating genetic and environmental data could elucidate growth dysregulation mechanisms [13].

Divergent structure is consistently found in frontotemporal, frontoparietal, limbic, and midline regions [14] implicated in social, emotional, and behavioral functions affected

in ASD [4,5,9]. But structural abnormalities are not ASD-specific, highlighting the need to model brain-behavior relationships [5].

Atypical cortical folding, influenced by early overgrowth, is also observed [4,5,6,9]. Functional MRI reveals aligned differences in activation and connectivity [15-17], though findings lack replication due to small samples and methodological inconsistencies [5,18,19].

On this atypical neurodevelopmental trajectory, questions arise:

- Can longitudinal neuroimaging data precisely characterize the timeline of early overgrowth and later accelerated decline?
- How do genetic and environmental factors interact to dysregulate developmental growth patterns in ASD?
- Do the regions showing divergent structure and function align with behavioral symptoms and change over time?

1.2.2. Preprocessing methods

While neuroimaging has yielded clues about ASD's neurobiological roots, robust biomarkers remain elusive [1-3]. Methodological challenges exist, including small samples, cross-sectional designs, and developmental factors [4-6].

Additional issues arise in MRI data acquisition and analysis. Scan quality variation, especially from head motion [7-9], requires rigorous quality assessment. Large open datasets demand automated quality control, but standard methods are lacking [10,11].

Preprocessing techniques like template registration may introduce confounds by obscuring group differences or reducing reproducibility [12-14]. Harmonisation techniques developed in neurotypicals could have similar effects [15].

Overall, poor standardisation of quality control and preprocessing likely contributes to inconsistent findings [16-18]. Recent efforts like BIDS, MRIQC, and fMRIPrep aim to establish standards and enhance reproducibility [19-21].

Several questions remain to be explored:

- What quality control methods are optimal for large multisite datasets vs. smaller single site data?
- How can preprocessing avoid obscuring real group differences or reducing reproducibility?
- What harmonisation techniques are appropriate for studies of divergent clinical populations?
- Is it possible to build a better preprocessing pipeline on MRI data to be able to study more accurately the brain characteristics of ASD?

1.2.3. Machine Learning approaches

Machine learning offers multivariate analytical advantages over univariate techniques [1,2]. Autism researchers have capitalised on ML's predictive capacity to build diagnostic classifiers from MRI [3-11].

However, many early studies lacked independent validation due to small samples [12,13]. When properly validated, ML approaches achieve moderate prediction

accuracy of 65-75% for ASD classification [14-17]. A large multisite challenge further demonstrated 70-80% accuracy, though performance declined on novel sites [18].

ML has limitations. Large samples are required, but available data may still be insufficient given ASD's heterogeneity and confounds like comorbidities [19-21]. Multisite differences and derivative inputs (e.g. volumetrics) can further bias results [22,23].

Deep learning is being explored to mitigate challenges. DL can learn predictive features directly from minimally processed data, reducing confounds from preprocessing [24]. But DL has its own challenges including architecture optimization, generalizability, reproducibility, and computational demands [25-27].

1.2.4. Deep Learning approaches

DL is a machine learning approach that learns hierarchical, multi-scale representations from raw data [1]. By minimising preprocessing, DL can learn predictive features directly [1]. Various DL architectures exist including convolutional neural nets for images and recurrent nets for sequences [1,2].

DL has shown initial promise for MRI-based ASD prediction, achieving accuracies of 65-75% [3-8]. However, a recent challenge found DL models tended to overfit compared to ML approaches [9].

DL has similar data demands as ML, needing large samples to mitigate ASD's heterogeneity [10]. It also remains sensitive to input quality and preprocessing biases [1,11]. Architectural complexity introduces challenges like overfitting and

intensive computation [12]. Sharing code and parameters openly for reproducibility is difficult [13].

This PhD aims to address these limitations by designing DL pipelines leveraging large open datasets. Goals include boosting predictive performance, enhancing model interpretability, and promoting reproducible practices [14].

On DL compared to ML, it can be wondered:

- Can DL methods mitigate limitations of ML like sensitivity to derivatives and heterogeneity?
- What DL architectures are most appropriate for multimodal prediction from MRI data?
- How can DL model optimization balance predictive performance with generalizability?
- What strategies enable reproducible DL workflows for MRI analysis?
- Can interpretability methods elucidate DL predictions and relate neural patterns to behaviour?

Given these initial results in neuroimaging, it appeared reasonable to hypothesise that:

- 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.

1.3. PhD project

1.3.1. Importance of the proposed research

Developing interpretable DL models that identify ASD neuroimaging biomarkers could advance precision psychiatry. Such tools could aid diagnosis, inform individualised interventions, and elucidate neural-behavioural links [1-3]. This could benefit clinicians, educators, families, and autistic individuals themselves [4,5].

For example, linking neurobiology to autistic traits could improve societal empathy and reduce discrimination in social and vocational settings [4,5]. Characterising early neural patterns may enable earlier intervention and improved outcomes [6,7]. Models could also track brain changes during treatment [8].

More broadly, this work aligns with evolving efforts to integrate neuroscience, AI, and genomics to better characterise mental health conditions based on underlying mechanisms [9]. Advanced analytics hold promise for precision medicine but require continued advances in techniques like interpretability to be clinically applicable [10,11].

1.3.2. 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 DL 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 DL pipelines to each MRI modality;
- Boost the acceptability of DL applications in medicine by improving the explainability and interpretability of models. This was achieved by finding and implementing methods to explain DL 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, DL, communication, leadership, management, and, more globally, in topics related to technology and Health.

1.3.3. Plan of the thesis/ Introduction of the studies

The work performed for this thesis is described 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 DL. 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 DL. 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 X. This project aimed to build a new DL 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.