The Brain Age Gap in Social Anxiety Disorder

by

Kimberly Vanessa Blake

BLKKIM003

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In fulfilment of the requirements for the degree

MSc (Med) Neuroscience

Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

Date of submission	26 July 2021
Supervisor	Dr Nynke Groenewold, Dep. of Psychiatry and Mental Health, University of Cape Town
Co-Supervisors	Dr Jonathan Ipser, Dep. of Psychiatry and Mental Health, University of Cape Town Prof. Dan Stein, Dep. of Psychiatry and Mental Health, University of Cape Town

DECLARATION

I, **Kimberly Vanessa Blake**, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: …KBlake……

Date: …26-07-2021…….

Abstract

Background: When an individual's brain appears 'older' than expected based upon their chronological age, they may be at an increased risk for developing brain-related diseases and cognitive decline. There is growing evidence of advanced brain ageing in neuropsychiatric diseases. Social anxiety disorder (SAD) is a disabling mental illness, which has been associated with both structural brain deficits and advanced biological ageing. However, brain age research has yet to be conducted in adults diagnosed with SAD. The present study investigated whether adults with SAD showed an advanced brain ageing process, compared to healthy controls (HCs), and whether brain ageing in SAD patients is associated with clinical characteristics.

Method: A systematic review of the literature was conducted to identify knowledge gaps in brain age research in psychiatric disorders before commencing with the present dissertation. Hereafter, a secondary data analysis of a large multi-site dataset was performed. T1-weighted structural MRI scans of 387 participants (SAD *n*=174, HC *n*=213) between the ages 18 and 60 years were included. These structural scans were segmented using both FreeSurfer and SPM12, after which they underwent quality control procedures. Brain age was estimated by two different machine learning models – Tobias Kaufmann's brain age model and James Cole's BrainageR. The primary outcome for analysis was the brain age gap (BAG), calculated by subtracting a participants' chronological age from their estimated brain age. General linear models were run to determine whether there was a significantly larger positive BAG in the SAD group (Kaufmann model *n*=100, Cole model *n*=155) compared to the HC group (Kaufmann model *n*=138, Cole model *n*=197) after adjusting for age, mean centred age^2 and sex. The association between BAG and comorbid depression and anxiety, as well as medication use and symptom severity, was also assessed.

Results: In the present study sample, predicted age was more strongly associated with chronological age for the Cole model estimates than the Kaufmann model estimates (Cole: Pearson correlation = 0.828, MAE = 4.78, *SD* = 3.96, versus Kaufmann: Pearson correlation = 0.576, MAE = 11.93, *SD* = 6.93). With the Kaufmann model, the SAD group had a significantly larger BAG than the HC group of almost one year (mean difference = 0.943 year, *SE* = 0.40, *p* = .019). In addition, with the Kaufmann model, patients without psychiatric comorbidities had a significantly larger BAG than HCs, of more than one year (mean difference = 1.242 year, *SE* = 0.49, *p* = .038). No difference was observed in BAG between patients with comorbidities and HCs (mean difference = 0.983 year, *SE* = 0.85, *p* = .749). In contrast, with the Cole model, the SAD group did not have a significantly larger BAG than the HC group (mean difference = 0.513 year, *SE* = 0.49, *p* = .383). Moreover, the Cole model found no significant difference in BAG between SAD patients with and without comorbidities, or between each of these groups and HCs (all *p* > .708). Finally, no significant associations were observed between the BAG and symptom severity and the BAG and medication use in SAD patients in the Cole or Kaufmann models.

Conclusion: This study observed contradictory evidence for a larger BAG between patients with SAD than HCs. The differences observed between the Cole model and the Kaufmann model may be a result of the different information used to estimate brain age (voxel-based volumetric data, compared to cortical thickness/surface area and subcortical/cerebellar volumes, respectively). The models demonstrated largely overlapping confidence intervals for group mean difference in BAG, suggesting that if there is a positive BAG in adults diagnosed with SAD, it is likely to be small. This should be verified in future research by using multiple different machine learning models based on different feature sets, to obtain more reliable and robust brain age estimates.

Acknowledgements

First, I would like to thank my supervisor, Dr Nynke Groenewold, for the countless hours spent guiding me through my project. Without her support and encouragement this dissertation would not have been possible. I would also like to thank my co-supervisors, Dr Jonathan Ipser and Prof Dan Stein. Dr Jonathan Ipser provided valuable feedback on the dissertation as well as essential help in the running of the brain age models. Prof Stein supported me throughout this process and provided a helpful round of feedback on the final draft of the dissertation.

I would also like to acknowledge the study participants and those who granted access to participant data. Thank you to Dr Janna-Marie Bas-Hoogendam and Prof Nic J.A. van der Wee for access to the mega-analysis data (Bas-Hoogendam et al., 2017). Thank you to the study primary investigators, Prof T. Straube, Prof K. Roelofs, Prof Nic J.A. van der Wee, Prof Dan Stein, Prof T. Furmark and Prof K. Luan Phan, for granting access to the data for inclusion in my dissertation.

Next, I would like to acknowledge the University of Cape Town's ICTS High Performance Computing Team: *hpc.uct.ac.za*, whose resources were used to run brain age model scripts.

Finally, I would like to thank Prof Dan Stein and the University of Cape Town for providing funding towards the completion of my degree. I am very grateful for the support.

Table of Contents

List of Figures

List of Tables

List of Supplementary Figures

List of Supplementary Tables

List of Abbreviations

ACC	Anterior cingulate cortex
APA	American Psychiatric Association
BAG	Brain age gap
BDI	Beck Depression Inventory
CIDI	Composite International Diagnostic Interview
CHR-E	Early clinical high-risk state of psychosis
CHR-L	Late clinical high-risk state of psychosis
CSF	Cerebral spinal fluid
DLPFC	Dorsolateral prefrontal cortex
DSM	Diagnostic and Statistical Manual of Mental Disorders
DTI	Diffusion tensor imaging
ENIGMA	Enhancing NeuroImaging Genetics through Meta-Analysis
FA	Fractional anisotropy
FEP	First episode of psychosis
GLMs	General linear models
GM	Grey matter
GMV	Grey matter volume
HCs	Healthy controls
HR-well	High-risk of developing a mood disorder who remained well
HR-MD	High-risk of developing a mood disorder who developed a mood disorder
ICD	International Classification of Disease
LSAS	Liebowitz Social Anxiety Scale
MAE	Mean absolute error

MDD	Major depressive disorder
MRI	Magnetic resonance imaging
NESDA	Netherlands Study of Depression and Anxiety
NCS-R	National Comorbidity Survey Replication
NCS-A	National Comorbidity Survey Adolescent
OFC	Orbitofrontal cortex
PANSS	Positive and Negative Symptom Scale
PICO	Population, influence, comparison and outcome
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
QC	Quality control
ROI	Region of interest
RVR	Relevance vector regression
SAD	Social anxiety disorder
SCID	Structured Clinical Interview for DSM-IV
SPM	Statistical parametric mapping
SVR	Support vector regression
UCT	University of Cape Town
VBM	Voxel-based morphometry
v-SVR	Linear v-support vector regression
WM	White matter

The Brain Age Gap in Social Anxiety Disorder

Introduction

Mental disorders, such as anxiety disorders, are major contributors to the global burden of disease (Rhem & Shield, 2019). Social anxiety disorder (SAD) is a common anxiety disorder (Stein & Stein, 2008), with lifetime duration of SAD affecting approximately 12.1% of the population in nationally representative surveys (Ruscio et al., 2008). SAD is a debilitating mental illness defined by excessive fear or anxiety of one or more social situations. This fear or anxiety causes clinically significant impairment in multiple aspects of an individual's life, such as in social or occupation settings (American Psychiatric Association, 2013). Based on data from the World Mental Health Initiative Survey, upper middle-income countries, such as South Africa, have the highest SAD-related impairment in role functioning. In addition, upper middle-income countries have the highest SAD persistence, as measured by the ratio of 12-month disorder prevalence to lifetime disorder prevalence (Stein et al., 2017).

In addition to impairment in social domains, there is strong evidence in support of structural and functional brain abnormalities in patients diagnosed with SAD (Bas-Hoogendam et al., 2017; Frick et al., 2013; Irle et al., 2010; Syal et al., 2012). Anxiety disorders have been linked with declines in cognitive performance and physical alterations of the brain, such as decreased grey matter density. Both of these findings have been observed during biological ageing, suggesting an advanced ageing process in anxiety disorders (Perna et al., 2016). Moreover, evidence suggests that anxiety disorders significantly increase risk of mortality (Meier et al., 2016).

As science develops, new and exciting methods become available to research the underlying mechanisms of brain-related disorders, such as SAD, and the brain changes associated with biological ageing. One such method is brain age prediction. Brain age

prediction utilises computational models that are trained on a large sample of healthy individuals, finding and learning common patterns from structural magnetic resonance imaging (MRI) data that correspond to chronological age. These "machine learning" models are then applied to the MRI scans of individuals who were not included in the training phase, to make an age prediction (Cole & Franke, 2017). The difference between a person's predicted brain age and their chronological age is called the "brain age gap" (BAG). A positive BAG occurs when a person's brain appears older than their chronological age, which may indicate underlying health problems or an increased risk of developing brain-related diseases (Cole & Franke, 2017).

While brain age research is new, it may have multiple important applications in the medical and mental health fields in the future. Brain age prediction has the potential to separate age-related structural and functional brain changes from disease-related brain changes in corresponding domains. Moreover, this method may be able to predict individualised brain ageing trajectories and determine reference curves for healthy brain ageing. In turn, this could allow early identification of unhealthy and harmful brain ageing when an individual's brain age deviates from their reference curve. From a clinical perspective, this could help with understanding disease progression in the individual and allow for tailored treatment interventions (Franke & Gaser, 2019). Much research is needed to further develop the brain age methodology and increase our understanding of brain ageing. Moreover, it is important that brain age research is conducted on prevalent brain-related disorders, such as SAD (Stein et al., 2017).

The following dissertation uses two novel machine learning models that predict an individual's age using information from structural T1-weighted 3T MRI scans of SAD patients and healthy controls (HCs). These MRI scans have been acquired from a large, multi-site dataset. Originally, a brain age model developed by Tobias Kaufmann (Kaufmann

et al., 2019), which uses cortical thickness and surface area data as well as subcortical and cerebellar volumes to estimate brain age, was chosen as the model to be used in this study. However, after it was determined that the model fit statistics were sub-optimal for the study data, a second brain age model developed by James Cole (Cole et al., 2017; Cole et al., 2015; Cole & Ritchie et al., 2018), which uses voxel-based morphometry data, was chosen to be used in addition to the Kaufmann model. Using the results from these models, I will investigate whether there is a possible advanced brain ageing process in adults diagnosed with SAD as compared to a group of HCs, indicated by means of a positive BAG. To the best of my knowledge, the BAG has not yet been explored in adults with SAD. The following dissertation will encompass four chapters, as follows:

The first chapter will contain a literature review in which I will start by defining SAD and its epidemiological findings. This will then lead to an outline of cognitive neuroscientific perspectives of SAD and the brain, where I discuss evidence of advanced biological ageing in SAD patients as shown by leukocyte telomere research and changes in brain structure, function and cognitive performance associated with ageing. Next, I describe structural deficits observed in SAD.

In addition to a literature review, chapter 1 contains a systematic review of the literature, titled: Brain Age in Psychiatric Disorders: A Systematic Review. This systematic review aims to determine whether there is an advanced brain ageing process observed in adults diagnosed with axis 1 psychiatric disorders. In addition, it aims to establish in which specific disorders the brain ageing process may be the most apparent. Mood disorders, psychotic disorders, and substance use disorders are all explored here. Finally, chapter 1 concludes with the aims and hypotheses of the dissertation.

Chapter 2 outlines the methods used. This dissertation uses structural T1-weighted 3T MRI scans that were previously collected at 10 study sites around the world. The study

design of the dissertation is described in this chapter, in addition to participant recruitment, clinical measures, scan parameters and the characteristics of the two brain age models. I describe the quality control procedures that were followed for the two brain age methods and the demographic information of the final participant samples. Next, I provide an overview of the statistical analyses to be used in the dissertation, including the project's main analysis, sub-analyses, and post hoc tests. The chapter concludes with an ethics statement.

Chapter 3 presents the statistical results, including model fit statistics and general linear models run to test the hypothesis that a positive BAG will be present in adults diagnosed with SAD as compared to HCs. Additional general linear models were run to explore whether the BAG is larger in SAD patients with or without psychiatric comorbidities and HCs, and to determine if there is an association between the BAG and symptom severity and medication use in SAD patients.

In the final chapter, I place my statistical findings in the context of the literature discussed in chapter 1. The size of the BAGs found in the present dissertation are considered alongside the size of the BAGs observed in other psychiatric disorders that are included in the systematic review. Associations between the BAG and clinical information are discussed in conjunction with past research on SAD and other mental disorders covered in the systematic review. I conclude my dissertation by discussing study strengths and limitations and by making recommendations for future research in brain ageing in SAD.

1. Literature Review

1.1. Social Anxiety Disorder: A Definition and Epidemiological Findings

Mental illnesses, such as anxiety disorders, are a major contributing factor to the health of individuals and populations (Jacob & Coetzee, 2018). Mental disorders cause approximately 7% of global disease burden and are especially prominent in high and upper-middle income countries (Rhem & Shield, 2019). Social anxiety disorder (SAD), a debilitating mental illness, is classified in the 5th version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) as a marked fear or anxiety of one or more types of social situations (American Psychiatric Association [APA], 2013). SAD is known to have an early age of onset, often arising during adolescence. The median age of onset varies between countries, ranging from 11 as observed in Poland to 26 in South Africa (Stein et al., 2017).

SAD is a common anxiety disorder (Stein & Stein, 2008) with a 12-month prevalence estimate for DSM-IV diagnosed SAD of 7.1% and a lifetime prevalence estimate of 12.1% (Kessler et al., 2005; Ruscio et al., 2008). Higher prevalence rates are seen in females than in males (Fehm et al., 2005) particularly during adolescence and young adulthood (Wittchen et al., 1999). Twelve-month prevalence rates tend to decrease with age, with the highest rates being observed in adolescents and the lowest in older adults (Kessler et al., 2012). However, according to the National Comorbidity Survey Replication (NCS-R; Kessler & Merikangas, 2004) and the Adolescent Supplement (NCS-A; Merikangas et al., 2009), lifetime prevalence is higher among adults (in the 18-64-year age group, 13%) than in adolescents (13-17-year-old age group, 8.6%; Kessler et al., 2012).

To make a diagnosis of SAD, International Classification of Disease (ICD) criteria (World Health Organisation, 1993), which is used primarily in European and clinical settings, or DSM criteria, which was used by most of the research sites in the present dissertation, are applied. The ICD-10 has similar diagnostic criteria to the DSM, requiring a marked fear and

avoidance of social situations, and significant emotional distress to make a diagnosis. In addition, the individual must experience one or more of the following: shaking or blushing, fear of vomiting, and fear or urgency of urination or defecation (World Health Organisation, 1993).

Both the DSM-IV and DSM-5 require an individual to experience marked fear or anxiety of one or more social situations, such as interactions with others (e.g. conversations), observed by others (e.g. while eating), and performance situations (e.g. speech giving). In addition, individuals fear showing anxiety symptoms that will be negatively evaluated (i.e. will be humiliating). The social situation(s) almost always provoke fear or anxiety and are avoided, causing clinically significant distress and impairment in social, occupational, or other important areas of functioning (APA, 2013).

However, there are a few notable differences between the DSM-IV and the DSM-5 (APA, 2013). The DSM-IV requires that individuals over the age of 18 are able to realize that their anxiety is unreasonable or excessive, whereas this is no longer a requirement in the DSM-5. Instead, the anxiety must be disproportionate to the threat of the feared situation after cultural factors have been taken into consideration. Finally, the DSM-5 now requires individuals of all ages to have a minimum of 6-months of symptom duration, and the specifier criterion has changed from 'specify if generalised' to 'specify if performance specific.' This last criterion can be used in the assessment of the severity of SAD. There is evidence in support of two sub-types of SAD – generalised SAD and SAD specific to performance fears. Generalised SAD usually describes a more severe and disabling form of the disorder due to the greater number of feared social situations. This subtype is also more persistent and more comorbid with other psychiatric disorders (Kessler et al., 1998).

Individuals diagnosed with SAD are at high risk for developing certain comorbid psychopathologies (Stein & Stein, 2008; Stein et al., 2017). Based on data from the World

Mental Health Survey Initiative, other anxiety disorders are the most common comorbidities in SAD (individuals with lifetime SAD and lifetime comorbid anxiety disorder: 59.8%), followed by mood disorders (individuals with lifetime SAD and lifetime mood disorder: 47%; Stein et al., 2017). Other less common comorbidities include substance use disorders and impulse control disorders (Stein et al., 2017). In females, the number of social fears is generally greater and comorbidities with other anxiety, depressive and bipolar disorders are more common. On the other hand, males with SAD commonly fear dating and comorbidities with oppositional defiant disorder and conduct disorder are more common, in addition to the use of substances to relieve disorder symptoms (Ruscio et al., 2008; Turk et al., 1998). Of note, SAD usually precedes comorbid psychopathologies (Beesdo et al., 2007).

1.2. SAD and the Brain

1.2.1. SAD and the Brain: Advanced Biological Ageing

The following section will consider different studies discussing advanced biological ageing in psychopathology. There are some indications that anxiety disorders, including SAD, may be associated with advanced biological ageing. Leukocyte telomere length is a well-known biomarker of ageing and has been found to be significantly shortened in patients with anxiety disorders compared to healthy controls (HCs) (Hoen et al., 2013; Malouff & Schutte, 2017; Verhoeven et al., 2016). Telomeres are specialised nucleic acid-protein complexes which protect DNA from damage by capping the ends of linear DNA. Leukocyte cells are commonly used for the measurement of telomere length. Each time a cell replicates, the ends of the telomere are not replicated, causing a progressive shortening (Verhoeven et al., 2016).

Further, anxiety disorders have been linked with poorer cognitive performance, such as poor attention, decreased grey matter (GM) density, frontal white matter (WM) alterations, and impaired functional connectivity of brain networks, which may be indicators of advanced

ageing (Baur et al., 2011; Perna et al., 2016). Similarly, a cohort study found that participants with a higher p-factor (as measured by persistence and variety of psychopathology) exhibited advanced biological ageing including hearing difficulties, visual difficulties, motor difficulties, balance difficulties and cognitive difficulties (Wertz et al., 2021). Examples of psychopathology as measured by the p-factor included SAD, MDD, schizophrenia and eating disorders. The signs of ageing observed in participants occurred between the ages of 25-45, years before the symptoms of ageing related diseases typically begin. Finally, Meier et al. (2016) conducted a population-based study using nationwide Danish register data and found that anxiety disorders significantly increased mortality risk from both natural causes (e.g. respiratory disease) and unnatural causes (e.g. suicide), especially when comorbid with depressive disorders.

The following section will expand on the brain-related deficits observed in SAD by discussing literature on structural magnetic resonance imaging (MRI) used to investigate grey matter volume (GMV) alterations and cortical thickness. GMV and cortical thickness will be explored to provide background and to motivate the use of voxel-based volumetric data and cortical thickness/surface area data and subcortical/cerebellar volumetrics in analyses in the present dissertation.

1.2.2. SAD and the Brain: Structural Deficits Observed in MRI Studies

Wang et al. (2018) conducted a meta-analysis of 11 studies using a whole-brain voxel-based morphometry (VBM) approach to explore GMV alterations in 470 SAD patients and 522 HCs. Larger GMVs were observed in the left precuneus, the right supplementary motor area, and the middle occipital gryus in SAD patients compared to HCs, whereas smaller GMVs were found in the left putamen. GMVs were larger in SAD patients who did not have a MDD comorbidity in the left superior parietal gyrus, extending to the precuneus, and in the right inferior temporal gyrus, extending to the fusiform gyrus (amongst other

regions). These patients had smaller GMVs in the bilateral thalami. SAD patients not currently taking medication had smaller GMVs in the left thalamus with no larger GMVs found (Wang et al., 2018). Finally, the study found similar results when looking at adult patients only, with no significant correlation with age. However, a commentary by Bas-Hoogendam (2019) notes that Wang et al. (2018) misrepresent details of studies included in the meta-analysis. Certain studies may have been erroneously excluded from a sub-group analysis which explored the effect of age on GMV alterations in SAD. Accordingly, Bas-Hoogendam (2019) warns that readers should interpret the results of this study with caution due to several methodological weaknesses.

Furthermore, an international multi-centre mega-analysis conducted earlier by Bas-Hoogendam et al. (2017) found contrasting results to those discussed above. The study used a region of interest (ROI) approach and VBM to compare the GMV of SAD patients (n=174) and HCs (n=213), finding no diagnosis-specific GMV differences in the prefrontal or parietal cortex. Larger GMV was found in the right putamen in SAD patients compared to HCs, and a positive correlation between the right putamen and symptom severity as shown by the total score of the Liebowitz Social Anxiety Scale (LSAS) was observed within the SAD group (Spearman's rho = 0.21, $p < .05$). No GMV alterations were found in amygdala-hippocampal structures. Lastly, there was no significant diagnosis-age interaction at the whole-brain level.

Of note, smaller scale research studies have reported enlarged bilateral amygdala and left hippocampal volume in SAD (Machado-de-Sousa et al., 2014) and smaller amygdala and hippocampal volumes in generalised social anxiety (n=24) (Irle et al., 2010). Here, smaller right hippocampal volume was significantly related to higher social anxiety severity (LSAS) in patients and right amygdala volume was significantly negatively predicted by age, indicating reduced amygdala volume in older compared to younger SAD patients (Irle et al., 2010). Syal et al. (2012) did not find significant differences in hippocampal and amygdala

volume between SAD patients and HCs. However, uncorrected right amygdala was smaller in the SAD cohort. Additional findings show GM deficits in the right amygdala in drug-naïve SAD patients (*n*=20) that are negatively correlated with SAD duration and positively correlated with age of onset (Meng et al., 2013). This suggests that an early onset and disorder progression may be associated with atrophy in the right amygdala (Meng et al., 2013). In addition, decreased GM density in the bilateral thalami, the right precuneus, and the right thalamus, have been reported in drug-naïve SAD patients (Meng et al., 2013).

Evidence suggests that certain structural anomalies are SAD specific. A study by Talati et al. (2013) found greater GMV in the left parahippocampal cortex and the cerebellum, and lower GMV in the temporal pole in SAD patients (*n*=16) compared to HCs (*n*=20). Interestingly, similar findings were seen in an independent clinical sample using the LSAS, and in a combined dataset of the two samples. These patterns were not seen in a panic disorder group, suggesting specificity of these findings to SAD (Talati et al., 2013).

Finally, other structural findings in SAD literature include increased cortical thickness in the right parietal cortex and dorsolateral prefrontal cortex (DLPFC; Brühl et al., 2014; Zhao et al., 2017). However, Irle et al. (2014) observed smaller GMV in the left inferior parietal cortex, with GMV density significantly negatively associated with avoidance symptoms. Syal et al. (2012) found cortical thinning in the right rostral middle frontal gyrus within the DLPFC in patients. Increased cortical thickness has been observed in the bilateral medial PFC, specifically the medial superior frontal cortex (Zhao et al., 2017), and the right anterior cingulate cortex (ACC; Brühl et al., 2014). One study reports a significant negative correlation between symptom severity (LSAS) and cortical thickness in the right rostral ACC (Frick et al., 2013). Finally, thinner cortical thickness in the right precentral cortex has been seen in SAD patients (Syal et al., 2012; Zhao et al., 2017). Zhao et al.'s (2017) findings were

also observed in their MDD group, suggesting possible common structural alterations between the two disorders.

In summary, structural alterations observed in SAD research vary with inconsistent results. Smaller studies report increased volume in the bilateral amygdala and left hippocampus (Machado-de-Sousa et al., 2014) and decreased volume in the amygdala and hippocampus in SAD patients (Irle et al., 2010). In contrast, the mega-analysis conducted by Bas-Hoogendam et al. (2017) did not find any structural alterations here, nor did they observe diagnosis specific GMV differences in the prefrontal or parietal cortex. Instead, the right putamen was implicated with larger GMVs found in SAD patients, while in contrast to this, Wang et al. (2018) found smaller GMVs in the left putamen of SAD patients compared to HCs. More research is needed to investigate these potential deficits further. These GMV and cortical thickness findings are relevant to this thesis, given that they form the basis of the machine learning based brain age models that have been applied to the study data.

1.2.3. SAD and the Brain: Machine Learning as a Tool for Diagnostic Classification

Machine learning is an artificial intelligence method which uses statistical analysis to learn patterns and make accurate predictions on unseen data (Cole & Franke, 2017). A study by Frick et al. (2014) aimed to discriminate between SAD patients and HCs by using a support vector machine model trained on GMV data. The method was able to successfully classify participants as SAD and HCs, with a balanced accuracy of 84.5% when using information from the whole brain. This provides evidence that the weights from support vector machine classifiers may have the potential to identify diagnostic brain-based biomarkers. However, the study used a small, male only sample in the test and training sets, thus findings should be interpreted with caution. Similarly, a study by Boeke et al. (2020) applied a stacked machine learning model, with 3 ridge regression base models, to predict trait anxiety from volume and cortical thickness data, region-to-region functional

connectivity and amygdala seed-to-voxel connectivity. Trait anxiety was significantly predicted in the model training dataset, however not in the unseen sample. Although findings were not in support of a generalisable anxiety biomarker, future research should explore this topic further due to the potential of biomarkers to advance psychiatry.

In addition to diagnostic classification, increasing evidence shows that machine learning methods can also be used to predict the chronological age of an individual by using structural MRI brain scans (Cole et al., 2017). The degree to which a person deviates from the 'normal' trajectory of ageing may indicate underlying health problems and 'older' appearing brains may be at an increased risk for developing brain-related diseases (Cole & Franke, 2017). This leads me to the following systematic review, in which the brain age estimation method is explored in psychiatric disorders.

1.3. Brain Age in Psychiatric Disorders: A Systematic Review

1.3.1. Rationale and Objectives

The brain is affected by ageing in numerous ways, including the alteration of molecular and cellular functioning and the biophysical structure of the brain (Cole, 2017). As the human brain ages, cognitive functions decline, neurodegenerative disease risk is increased and symptom severity of neuropsychiatric diseases intensifies (Cole, Marioni, et al., 2018). Evidence suggests that certain diseases, such as psychiatric disorders, advance the ageing of the brain (Cole et al., 2019), potentially contributing to this deviation from the norm.

A promising method to investigate the association between the ageing brain and disease is "Brain Age Gap Estimation", which uses data-driven learning methods to make age predictions at the individual or group level (Franke & Gaser, 2019). Brain age models are trained on large datasets of HCs, learning patterns corresponding to chronological age from structural or functional MRI scans. Structural MRI was chosen for inclusion in the systematic review due to evidence of structural brain changes observed in ageing (Cole, 2017) and because

structural MRI has been used in the present dissertation. Multivariate machine learning regression models (e.g. Gaussian processes regression, support vector regression [SVR], relevance vector regression [RVR]) are developed to capture these patterns in a training dataset and are subsequently used to predict the chronological age (brain age estimation) of individuals who were not included in the training set (Cole & Franke, 2017). In addition, deep learning methods (e.g. convoluted neural networks) that use artificial neural networks to make brain age predictions are becoming increasing popular in brain age research (Cole et al., 2017). The deviation between a person's chronological age and their estimated brain age is termed the "brain age gap" (BAG). When an individual's brain appears to be 'older' than their chronological age, this is referred to as a positive BAG (Cole, Marioni, et al., 2018). This method has considerable potential as a biomarker of brain health (Cole & Franke, 2017).

An ageing biomarker provides a measurement of a person's age based on the biological age of a cell, tissue or organ (Cole & Franke, 2017). The BAG is considered a reliable marker of ageing with studies measuring accuracy through cross-validation and testing models in independent samples (Cole et al., 2019). A study conducted in 73-year-old participants found brain age to have a significant relationship with mortality risk and a higher brain age was associated with slower walking time, lower grip strength, lower forced expiratory volume, and lower fluid cognitive performance (Cole, Ritchie, et al., 2018). In addition, brain age was found to be a stronger predictor of mortality than other more established ageing biomarkers, such as leukocyte telomere length (Cole, Ritchie, et al., 2018). The following systematic review will help to better understand the relationship between the ageing brain and psychiatric disease by discussing articles which assess the BAG across psychiatric disorders at different disease stages, while simultaneously looking at a variety of clinical and sociodemographic variables, using varying forms of brain age methodology.

The following systematic review aimed to determine whether there is an advanced brain ageing process in adults diagnosed with an axis 1 psychiatric disorder. This was investigated by systematically assessing records which have explored the BAG estimation, based on structural MRI scans, in adults diagnosed with axis I psychiatric disorders as compared to HCs. Adult samples were chosen for this systematic review as the present dissertation included adult samples in analyses. Moreover, it is difficult to interpret the BAG during childhood and adolescence due to brain maturational processes. Several clinical factors are explored that may amplify the magnitude of the BAG, including type of diagnosis, medication use, stage of disease, and conversion from an at-risk mental state to a diagnosable disease.

1.3.2. Systematic Literature Search

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) was followed (Moher et al., 2009). Specific search terms were applied in PubMed and Scopus for the identification of relevant papers (see Table 1 for details). Records were exported from PubMed and Scopus and organised in a spreadsheet starting with oldest records to most recent. Publication dates of retrieved records ranged from 1986 to 2020. All records with a publication date before 2010 were excluded as the first brain age paper was published in 2010 (Franke et al., 2010). This was identified in the reference list of an article reviewing brain age research (Franke & Gaser, 2019). This resulted in the exclusion of 831 papers. Next, 84 duplicates were excluded, resulting in 1526 records for the title/abstract review.

Table 1.
Database Search Terms, Date, and Number of Hits

Database	**Search**	**No. of record hits**	**Date of search**
PubMed	mental disorders [mh] OR (depressive disorder, major [mh] OR depressive disorder [mh] OR major depressive disorder [tw] OR depressive disorder [tw]) OR (bipolar disorder [mh] OR bipolar disorder [tw]) OR (schizophrenia spectrum and other psychotic disorders [mh] OR psychosis [tw]) AND (brain [tw] OR brain [mh]) AND (aging, premature [mh] OR aging [mh] OR ageing [tw] OR aging [tw]) AND (magnetic resonance imaging [mh] OR MRI [tw]) *Note:* 'humans' filter was selected	2121	22/10/2019
Scopus	(("mental disorders" OR "major depressive disorder" OR "bipolar disorder" OR "psychosis") AND brain AND ("premature aging" OR "premature ageing" OR aging OR ageing) AND ("magnetic resonance imaging" OR mri))	320	22/10/2019

In review of titles and abstracts, records were evaluated by myself and an independent reviewer (AA) using a set of pre-defined inclusion and exclusion criteria following PICO guidelines (population, influence, comparison and outcome; Moher et al., 2009). Papers which passed these criteria, or for which there was insufficient information provided to determine whether they passed the criteria, were selected for full text review. The full text articles were reviewed using a more detailed set of inclusion and exclusion criteria, as certain criteria, such as the use of a brain-based marker of ageing as outcome, are easier to assess in full text where more information is provided. Where discrepancies between reviewers arose, the decision of inclusion or exclusion was made by a third independent reviewer (NG) in consensus with the initial reviewers. This resulted in 10 papers included for data extraction. For a detailed description of these criteria, please see Table 2 below.

Table 2.

Inclusion and Exclusion Criteria for Systematic Review Stage 1 and 2

PICO criteria	Title/Abstract review	Full-text review
Population	**Included:** adult human participants between the ages of 18-65	**Included:** adult population is used, with mean age of participants being no less than 18 years and no more than 65 years
Influence	**Included:** axis 1 psychiatric disorders **Excluded:** if disorder onset is a result of an organic cause; or if participants received any of the following as a primary diagnosis: neurodevelopmental disorders, disruptive disorder, impulse-control disorders, conduct disorders, and neurocognitive disorders	**Included:** axis 1 psychiatric disorders, diagnosed using DSM-IV, DSM-5 or ICD-10. The group must be used for application of the brain age estimation **Excluded:** if disorder onset is a result of an organic cause; or if participants received any of the following as a primary diagnosis: neurodevelopmental disorders, disruptive disorder, impulse-control disorders, conduct disorders, and neurocognitive disorders
Comparison	**Included:** psychiatric group is compared to a HC group	**Included:** psychiatric group is compared to a HC group
Outcome	**Included:** brain-based marker of ageing is used; with structural MRI input	**Included:** A machine learning algorithm (supervised or unsupervised) is used to determine the 'brain age' and calculate the BAG of the psychiatric and HC groups; the 'brain age' is based on a set of measurable features using thickness, surface area or volume, or diffusivity and anisotropy measures **Excluded:** if the brain age model has been trained on specific ROIs only
Other	**Included:** original article, written in English	**Included:** if sample size of psychiatric and HC groups is at least *n*=20, per group

Note. BAG, brain age gap; DSM, Diagnostic and Statistical Manual of Mental Disorders; HC, healthy control; ICD, International Classification of Diseases; MRI, magnetic resonance imaging; ROI, region of interest.

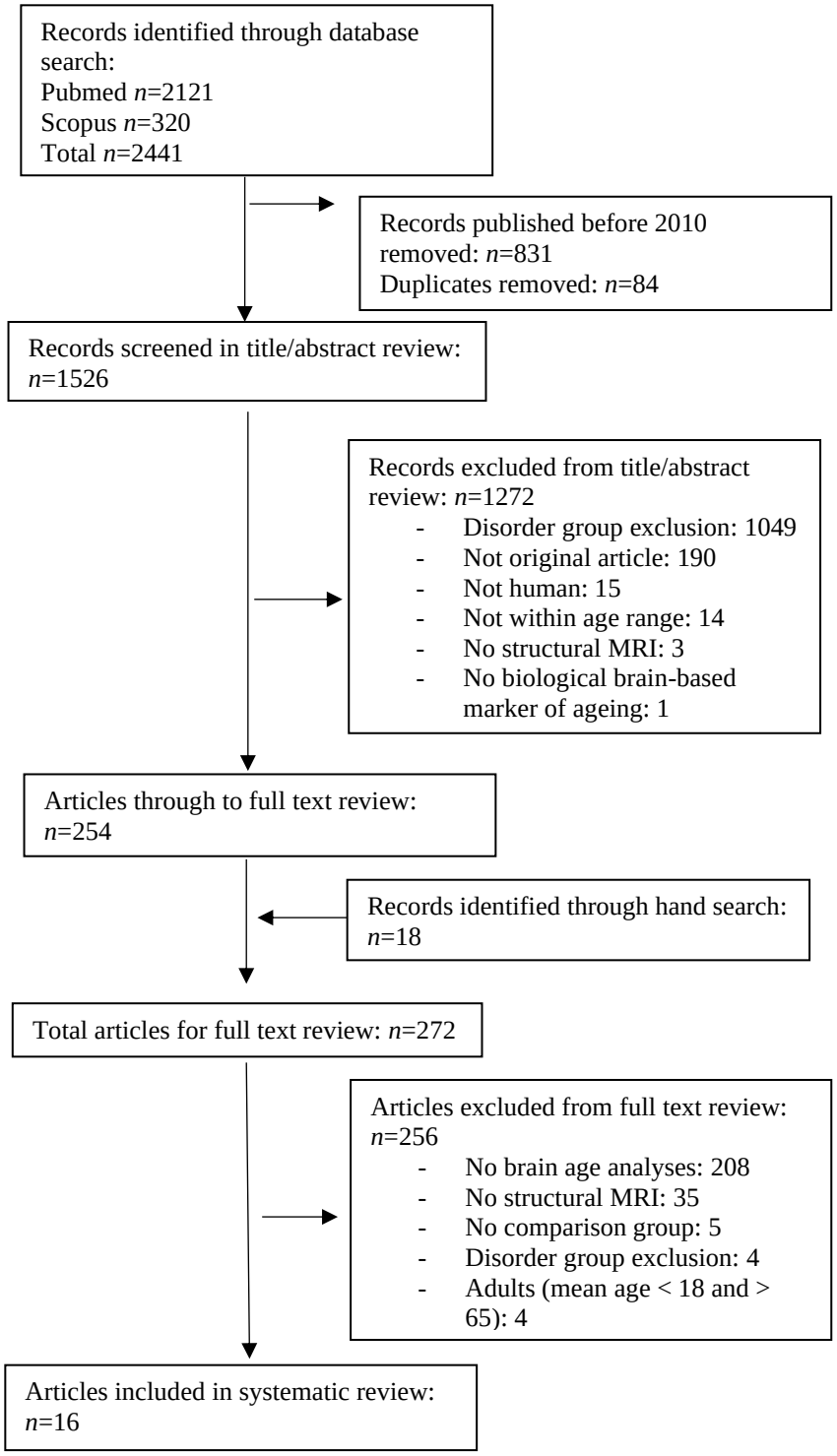
Reference lists of reviews were examined to ensure no relevant papers had been missed (Cole, Marioni, et al., 2018; Cole & Franke, 2017; Franke & Gaser, 2019). In addition, a hand search was conducted by checking unpublished papers in bio-archives (https://www.biorxiv.org/) as well as going through a list of papers that have cited the 10 eligible studies included in this systematic review. Seventeen extra papers from the citation

search and bio-archives were included for full text review. Of these, six were chosen for inclusion in the systematic review, three of which were awaiting publication. Papers awaiting publication were included in this review as the inclusion of grey literature may reduce publication bias (Paez, 2017). Finally, I contacted the corresponding authors of the included papers requesting any additional unpublished results that may be available for sharing. In total, 16 original articles were identified as eligible for inclusion.

Data was managed using Excel and Endnote. The following information was extracted: MRI modality, type of brain age model, brain age modality (GM, WM or both), mean absolute error (MAE), sample size, diagnostic classification, medication use, number of males and females, sample chronological age and estimated BAG for cases and controls. A systematic review was chosen rather than a meta-analysis, as multiple articles described partly overlapping psychiatric samples, resulting in a lack of sample independence. The extracted results were structured according to diagnostic categories and synthesised by comparing the major disorder groups and brain age methodology.

Figure 1.

PRISMA Flowchart: Record Screening Process



1.3.3. Summary of Search Results

The included articles investigated the BAG in psychotic disorders (*n*=3), mood disorders (*n*=4), psychotic disorders and mood disorders (*n*=7), with one additional paper exploring this method in alcohol dependence.

Table 3.

Sample and Brain age Model Information per Included Article

Article	Test Group	Total *n*	Female n (%)	*M* age (*SD*)	Medication use	Model	HC training *n*	MAE (yrs)	BAG Modality: GM and/or WM	BAG	Scan Modality
Koutsouleris et al., 2014	MDD	104	52 (50.00%)	42.30 (12.00)	AP: 17.3% AD: 73.1% Li: 6.7%	v-SVR	800	4.6	GM	+4 (6.2)	T1
	SCZ	141	33 (23.40%)	28.50 (7.32)	AP: 90% AD: 7.9% MS: 3.6%					+5.5 (6)	
	HC	437	214 (48.97%)	32.60 (10.90)	N/A					-	
Koutsouleris et al., 2015	MDD	104	50.00%	42.30 (12.00)	TAP: 10% AAP: 9% MS: 13% Li: 7%	v-SVR	800		GM	+4 (6.2)	T1
	SCZ	158	26.00%	30.80 (10.00)	TAP: 30.7% AAP: 67.3% AD: 7.9% MS: 3.3%					+6 (6)	
	BD	35	49.00%	39.00 (9.60)	AAP: 40% AD: 16.7% MS: 66.7% Li: 16.7%					+3.8 (6.5)	
	FEP	23	26.00%	26.80 (6.50)	AAP: 39.1% AD: 21.7%			4.6		+5.1 (8.5)	
	HC	437	214 (48.97%)	32.60 (10.90)	N/A					-	

Schnack et al., 2016	SCZ	Baseline: 341	[1]Not available	29.50 (9.96)	TAP: 30% AAP: 62% Medication naïve: 11%	v-SVR	386	4.31	GM	+3.36 (5.87)	T1
		Follow-up1: 192	[2]Not available	31.61 (9.58)						+4.32 (6.20)	
	HC	Baseline: 386	179 (46.37%)	34.07 (11.81)	N/A					-0.0017 (5.40)	
		Follow-up1: 186	[3]Not available	35.98 (12.99)						- 0.045 (5.82)	
Guggenmos et al., 2017	ALC	119	15.10%	45.00 (10.7)	None	MRR	97	6.9	GM	+4 (0.7)	T1
	HC	97	16.50%	43.70 (10.80)	N/A					0.0	
Nenadic et al., 2017	SCZ	45	16 (35.55%)	33.70 (10.50)	[4]AP: n = 30 (88.24%)	RVR	410	N/A	GM	Total: +2.56 (6) Males: +3.37 (6.75) Females: +1.07 (4.13)	T1
	BD	22	12 (54.55%)	37.70 (10.70)	All BD patients were on MS and/or AP medication					Total: -1.25 (4.63) Males: -1.85 (6.09) Females: -0.76 (3.17)	
	HC	70	30 (42.86%)	33.80 (9.40)						Total: -0.22 (5.57) Males: -0.93 (5.26) Females: +0.71 (5.90)	

Chung et al., 2018	CHC-NC all ages	236	94 (39.83%)	17.30 (2.10)	AP: 22.5%	PRR	953	Training HC: 1.69. Test HC: 1.41	GM	17-21yo: +0.58 (2.08)	T1
	CHC-C all ages	39	16 (41.03%)	17.20 (2.23)	AP: 30.8%					17-21yo: +0.06 (2.09)	
	HC all ages	109	43 (39.45%)	17.02 (2.44)	N/A					17-21yo: +0.11(1.62)	
Kolenic et al., 2018	FEP	120	46 (38.33%)	27.00 (4.94)	1st gen AP: 13.79% 2nd gen AP: 98.28%; AC: 4.31%; AD: 13.79%; Li: 0.86%	RVR	504	% error: 16.29%	GM and WM	+2.64 (1.80-3.48)	T1
	HC	114	61 (53.51%)	25.70 (4.01)	N/A					+0.14 (-0.73-1.00)	
Besteher et al., 2019	MDD	38	21 (55.26%)	45.65 (15.68)	PM >14 days: n = 27 PM < 14 days: n = 6	RVR	743	N/A	GM	+0.412	T1
	HC	40	20 (50%)	42.00 (13.17)	N/A					0.00	
De Nooij et al., 2019	HR-MD	TP1: 35	21 (60%)	TP1: 1.40 (5.20)	None	RVR	167	2.21	GM	TP1: +0.01 (1.39)	T1
		TP2: 31		TP2: 23.70 (5.00)	PM: n = 6					TP2: +0.30 (1.30)	
	HR-well	TP1: 74	36 (48.65%)	TP1: 21.60 (4.90)	N/A					TP1: +0.36 (1.22)	
		TP2: 47		TP2: 23.80 (5.20)	N/A					TP2: +0.08 (1.04)	

	HC	TP1: 93	51 (54.84%)	TP1: 21.60 (3.80)	N/A					TP1: +0.04 (1.14)	
		TP2: 46		TP2: 23.40 (3.60)	N/A					TP2: +0.36 (1.01)	
Hajek et al., 2019	FEP	43	17 (39.53%)	27.09 (4.93)	None	RVR	504	N/A		+2.64 (4.15)	T1
Study 1	HC	43	17 (39.53%)	27.05 (4.40)	N/A					-0.01 (4.15)	
Study 2	Unaffected HR	48	29 (60.42%)	20.91 (4.15)	N/A	RVR	504	N/A	GM and WM	-1.02 (5.02)	T1
	Affected familial	48	33 (68.75%)	23.09 (4.51)	AC = 5 AD: 11 AP: 9 Li: 2 Lifetime history Li: 7					-0.96 (5.18)	
	HC	60	36 (60.00%)	23.41 (2.93)	N/A					+0.25 (5.27)	
Kaufmann et al., 2019	MDD	208	-	Range: 18-71	N/A	xgboost	35474	N/A	GM and WM	+0.86	T1
	SCZ	1110	-	18-66	N/A					+3.83	
	PSYMIX	300	-	18-69	N/A					+1.49	
	BD	459	-	18-66	N/A					+2.07	
	HC	HC group per diagnostic group with same *n*	-	3-95						-	
Shahab et al., 2019	SSD	81	33 (40.74%)	Old: 64.18 (8.52) Young: 33.10 (8.58)	N/A	RF	50	N/A	GM and WM	+7.8 (1.62)	T1 DTI CP

Analysis1	BD	53	25 (47.17%)	O: 62.68 (8.02) Y: 27.10 (6.01)						+0.14 (2.33)	
	HC	41	-	O: 65.74 (8.99) Y: 27.25 (7.67)						delta and SE: +0.67 (± 2.30)	
Analysis2	SDD	67	-	See above		RF	50		GM and WM	+6.12 (1.99)	T1 DTI
	HC	30	-	See above						+1.80 (2.83)	
Tonneson et al., 2019	SCZ	648		34.49 (11.40)	N/A	xgboost	927	Training HC: 7.28. Test HC: 7.98	GM and WM	-0.73 (±8.79)	DTI
	BD	185		33.12 (10.53)	N/A					-1.00 (±8.37)	
	HC	990		34.70 (11.24)	N/A					-3.29 (±7.89)	
Van Gestel et al., 2019	BD Li	41	23 (56.10%)	46.96 (13.77)	AD: 39.02%	RVR	504	4.62	GM	mean [95% CI]: -0.52 [-2.31; 1.26]	T1
					AP: 34.15%				WM	–1.97 [–3.62; –0.32]	
					AC: 39.02%				GM and WM	+0.83 [–0.88; 2.54]	
	BD no Li	43	26 (60.47%)	48.17 (11.49)	AD: 48.84%					+4.29 [2.54; 6.04]	
					AP: 48.84%					–0.46 [–2.07; 1.16]	
					AC: 72.09%					+4.96 [3.29; 6.64]	

	HC	45	21 (46.67%)	42.26 (13.80)	N/A					-1.25 [–2.97; 0.48] -2.52 [–4.11; –0.93] -0.15 [–1.80; 1.50]	
Kambeitz-Ilankovic et al., 2019	CHR	48	16 (33.33%)		N/A	SVR	36	2.2	GM	+4.3 (8.1)	T1
	CHR-E	19	7 (36.84%)	25.20 (5.68)	N/A					+2.3 (6.3)	
	CHR-L	29	9 (31.03%)	24.35 (5.88)	N/A					+5.6 (9.0)	
	HC	36	17 (47.22%)	27.16 (3.71)	N/A					-0.1 (5.5)	
	Converters	15			N/A					+5.5 (4.5)	
	Non-converters	33			N/A					+4 (6.6)	
Han et al., 2020	MDD	2675	1689 (63.14%)	-	AD: n = 1717	SVR	2188	Training HC males: 6.32 (5.06); females: 6.59 (5.14). Test HC males: 6.50 (4.91); females: 6.84 (5.32)	GM	+1.08 (SE: .22)	T1
	HC	4314	2435 (56.44%)		N/A					-0.20 (8.44)	

Note. AC, anticonvulsants; AD, antidepressants; ALC, alcohol dependence; AP, antipsychotics; AAP, atypical antipsychotics; BD, bipolar disorder; CHR, clinically high-risk; CHR-E, clinically high-risk early, CHR-L, clinically high-risk late; CP, cognitive performance; DTI, diffusion tensor imaging; HC, healthy control; HR, high risk; Li, lithium; MAE, mean absolute error; MDD, major depressive disorder; MRR, multi-linear ridge regression; MS, mood stabiliser; PM, psychotropic medication; PRR,

penalised ridge regression; PSYMIX, mixed diagnosis in the psychosis spectrum; RF, random forest; SCZ, schizophrenia; SS, sample size; SSD, schizophrenia, schizoaffective disorder, schizophreniform disorder, or psychotic disorder not otherwise specified; TAP, typical antipsychotics. Affected familial participants refer to participants who have a parent diagnosed with bipolar disorder and have at least one mood disorder themselves. [1] Approximately 21.65% of participants were female, the exact numbers are not known. [2]Approximately 21.98% of participants were female, the exact numbers are not known. [3]Approximately 28.93% of participants were female, the exact numbers are not known. [4]This is an estimate based on a smaller sample (Nenadic et al., 2015).

1.3.4. Brain Age in Psychotic Disorders

Starting with the largest brain age study to date, Kaufmann et al. (2019) investigated the BAG in a variety of psychiatric disorders compared to HCs, consisting of multiple samples. Brain age was predicted using a gradient tree boosting (xgboost) machine learning model trained on cortical thickness, surface area and volumetric features of the raw T1 structural images of HCs. The model was trained separately on males and females (ages 3-89 years). A positive BAG was found in the schizophrenia group (*n*=1110, +3.83 years, *d* = 0.51) and the mixed diagnoses in the psychosis spectrum group (*n*=300, +1.49 years, *d* = 0.21) compared to HCs (*n*=1110 and *n*=300, respectively). Evidence for advanced brain ageing was also found by Koutsouleris et al. (2014), using a linear v-support vector regression (v-SVR) brain age model trained on the GMV of HCs. A positive BAG of +5.5 years was found in the schizophrenia group (*n*=141) compared to HCs (*n*=437). Using the same model, training sample, and a slightly bigger patient sample size, Koutsouleris et al. (2015) found a BAG of +6 years in schizophrenia (*n*=158) compared to HCs (*n*=437).

Using linear v-SVR trained on the GM density of HCs, Schnack et al. (2016) calculated the BAG at a baseline scan and a follow-up scan approximately three and a half years later. A significant positive BAG of +3.36 years was found in schizophrenia patients at the baseline scan (*n*=341), and an even larger significant BAG was seen at the follow-up scan (*n*=192, +4.32 years) compared to HCs (BAG at baseline: - 0.0017 years; BAG at follow-up: - 0.045 years), indicating an interaction between group and BAG over time. Nenadic et al. (2017) found a positive BAG in their schizophrenia group (*n*=45) of +2.56 years, compared to controls (*n*=70, -0.22 years) when using RVR trained on the GMV of HCs. A BAG of +3.37 (*n*=29) was found in male patients only and of +1.07 (*n*=16) in females only.

Tonneson et al. (2019) took a different approach to the papers discussed thus far by training six models using xgboost on multiple modalities. Their main model used fractional

anisotropy (FA), mean diffusivity, radial diffusivity and axial diffusivity obtained from HCs using diffusion tensor imaging (DTI). A BAG of -0.73 was found in schizophrenia patients (*n*=648) compared HCs (*n*=990, -3.29). Two analyses were conducted by Shahab et al. (2019), the first of which calculated the BAG on multimodal data using cortical thickness in 78 brain regions, FA in 38 brain regions, and nine cognitive performance scores, using a random forest algorithm of 10 000 trees. A large positive BAG of +7.8 years was seen in the schizophrenia spectrum group (*n*=81), compared to HCs (*n*=41) who had a BAG of +0.67 years. The analysis was repeated without the inclusion of cognitive performance scores. This resulted in a slightly smaller BAG of +6.12 years for the schizophrenia spectrum group (*n*=67) and a slightly bigger gap of +1.80 years for HCs (*n*=30).

Three studies focused on the early stages of schizophrenia by investigating the BAG in first episode of psychosis (FEP) patients (Hajek et al., 2019; Kolenic et al., 2018; Koutsouleris et al., 2015). Koutsouleris et al. (2015) found a BAG of +5.1 years in their FEP group (*n*=23). Hajek et al. (2019) and Kolenic et al. (2018) used the same brain age methodology applied to patients undergoing their first hospitalization. Hajek et al. (2019) found a positive BAG of +2.64 years (*n*=43), compared to HCs (*n*=43, -0.01 years). A post hoc analysis showed that FEP patients and HCs differed only in brain age scores estimated from GM and not WM. Using the same FEP sample as above, differing in sample size (*n*=120), Kolenic et al. (2018) found a positive BAG of +2.64 years compared to controls (*n*=114, +0.14 years). A relationship with body mass was detected, with the highest BAG found in participants with FEP and obesity/overweight (+3.83 years), and the lowest in normal weight HCs. This suggests that being overweight, which is common in psychotic disorders, and may represent an adverse effect of antipsychotic medications, could be a contributor to the neurostructural changes which are seen already in the early stages of the disease (Kolenic et al., 2018).

Chung et al. (2018) trained ridge regression models on HCs aged 3-21 years. Study participants were separated into clinically high-risk individuals who did not convert to psychosis (*n*=236), clinically high-risk individuals who converted to psychosis post-scan (*n*=39), and HCs (*n*=109). In a sample of 17-21-year-old individuals, no significant differences were observed in the BAG between the three groups (HC: *n*=67, +0.11(1.62); non-converters: *n*=120, +0.58 (2.08); converters: *n*=22, +0.06 (2.09)). In younger adolescents (< 17 years), the BAG was significantly greater than 0 for converters (*n*=17, +1.59 years) and non-converters (*n*=125, +0.67 years). The mean BAG was not significantly greater than 0 in HCs (*n*=67). Moreover, converters had a significantly bigger BAG than HCs. Therefore, a deviation in neuro-maturation linked with psychosis risk was only noted in young adolescence.

Kambeitz-Ilankovic et al. (2019) used an SVR model including 14 neurocognitive measures to predict brain age in individuals with an early stage of clinical high-risk state of psychosis (CHR-E) and a late ultra-high-risk state of psychosis (CHR-L). High risk individuals were identified using the Basel Screening Instrument for Psychosis (Riecher-Rössler et al., 2008; Koutouleris et al., 2012). Both the CHR converters to psychosis (*n*=15) and the non-converters (*n*=33) showed a significantly higher BAG than HCs (*n*=36; converters: +5.5 (4.5) years; non-converters: +4 (6.6) years). The highest BAG was observed in the CHR-L group (+5.6 (9.0) years) compared to HCs (-0.1 (5.5) years). The CHR-E group did not have a significantly different BAG than the HCs. Lastly, a negative relationship between chronological age and brain age was observed, especially in younger CHR participants. The brain age delay observed in younger participants appeared to stabilize in older CHR participants in their early thirties. This may suggest that the BAG represents a neurocognitive delay in psychosis, rather than static neurocognitive impairment.

1.3.5. Brain Age in Mood Disorders

Kaufmann et al. (2019) also investigated brain age in mood disorders. As opposed to the large positive BAGs seen in their psychotic disorder groups, a small positive BAG of +0.86 years was found in MDD participants (*n*=208; *d* = 0.10). Similarly, using a ridge regression model trained on 2188 HCs, Han et al. (2020) observed a small, uncorrected BAG of +0.68 (8.82) years in individuals with MDD (*n*=2675) compared to -0.20 (0.22) in HCs (*n*=4314). A BAG of +1.08 (SE: 0.22) years significantly higher than controls was found in the MDD group when adjusting for age, age^2, sex and scanning site. Due to the higher errors shown by the model in participants older than 60 years, a sensitivity analysis including only patients between the ages of 18-60 years was performed, resulting in a slightly increased difference in MDD (+1.16, *SD* = 0.24) compared to HCs.

A smaller study which investigated the BAG in MDD (*n*=38) found no significant difference in mean BAG between patients (0.41) and HCs (*n*=40; 0.00) (Besteher et al., 2019). Although this does not support advanced ageing in MDD as observable by the brain age approach, this finding could be due to the small sample size used, or due to not including patients with comorbidities when these are commonly seen in individuals with MDD. Finally, in contrast to the small BAGs observed in the articles discussed thus far, Koutsouleris et al. (2014) and Koutsouleris et al. (2015) found a substantial positive BAG of +4 years in their MDD group (*n*=104). There are a few possible explanations for the big difference in size of BAG found by Koutsouleris et al. (2014) and Koutsouleris et al. (2015) compared to the other studies, as suggested by Han et al. (2020). The Koutsouleris et al. (2014) and Koutsouleris et al. (2015) studies used high-dimensional whole-brain GM maps whereas Han et al. (2020) used a smaller number of input features using FreeSurfer ROIs. However, Kaufmann et al. (2019) included multimodal parcellations and found a smaller brain age difference in MDD. In addition, the Koutsouleris et al. (2014) and Koutsouleris et al. (2015) studies used a

smaller number of scanners, training sample size, and MDD test sample size than Han et al. (2020) and Kaufmann et al. (2019).

The BAG was also explored in bipolar disorder. Here, Koutsouleris et al. (2015) found a positive BAG of +3.8 years (*n*=35), which is slightly smaller than the BAG seen in their MDD group. However, Kaufmann et al. (2019) found a BAG of +2.07 years (d = 0.29) in their bipolar disorder group (*n*=459), this being bigger than the BAG seen in their MDD group. Unlike in their schizophrenia group, Nenadic et al. (2017) found no significant difference in BAG between bipolar disorder patients (*n*=22, -1.25) and HCs (*n*=70, -0.22, p = 0.34). This was still the case when assessing the BAG in males and females separately. Based on DTI data, Tonneson et al. (2019) found a BAG of -1.00 in bipolar disorder patients (*n*=185) as compared to HCs (*n*=990, $p < .001$). Of note, this negative BAG is slightly bigger than the one seen in the schizophrenia group. Finally, in contrast to the large positive BAG found in their schizophrenia spectrum group, Shahab et al. (2019) observed a small positive BAG of +0.14 years in bipolar disorder patients (*n*=53), based on T1, DTI and cognitive performance data. In addition, bipolar disorder patients and HCs were found to be statistically equivalent in their predicted age.

Van Gestel et al. (2019) investigated the effects of lithium treatment on brain age in bipolar disorder patients. RVR was trained on the whole brain (GM and WM) of HCs and a significantly greater BAG was found in non-lithium treated bipolar disorder patients (*n*=43, +4.96), compared to lithium treated bipolar disorder patients (*n*=41, +0.83), and HCs (*n*=45, –0.15). Findings were similar for brain age estimated from GM only (non-lithium bipolar disorder = +4.29 years; lithium treated bipolar disorder = -0.52 years; controls = -1.25 years), however were no longer significant for brain age estimated from WM only.

Using the same model as Van Gestel et al. (2019), Hajek et al. (2019) investigated the BAG in offspring of parents diagnosed with bipolar disorder. Participants were divided into

familial high-risk but unaffected (*n*=48, individuals with no lifetime history of a mood disorder), and familial affected (individuals who met the criteria of an axis 1 mood disorder, MDD = 26, bipolar disorder = 19, other = 3). Unaffected participants had a brain age score of −1.02 (*SD* = 5.02), affected participants of −0.96 (5.18), and controls (*n*=60) of +0.25 (5.27), with no significant findings. Post hoc analyses showed no significant differences between the groups in BAG scores estimated from GM or WM.

Similarly, de Nooij et al. (2019) calculated the BAG in participants with a high-risk of developing a mood disorder who remained well (HR-well, timepoint 1 *n*=74, timepoint 2 *n*=47), those who developed a mood disorder (HR-MD, timepoint 1 *n*=35, including 6 bipolar disorder, timepoint 2 *n*=31) and HCs (timepoint 1 *n*=93, timepoint 2 *n*=46), at two timepoints. High-risk was defined as having one or more first-degree relatives, or two or more second-degree relatives, diagnosed with bipolar disorder. The brain age model was trained on the GM of a comparatively small sample of HCs (*n*=167) using RVR. All participants remained well at the first timepoint and BAG did not differ significantly between the three groups. Two years later, the mean BAGs were +0.36 (SD = 1.01) for HCs, -0.08 (SD = 1.04) for HR-well, and -0.30 (SD = 1.30) for HR-MD. However, after multiple comparison correction the brain age scores were no longer significantly different between the HR-MD group and HR-well group, and the HR-MD group and HCs.

1.3.6. Brain Age in Substance Use Disorders

While most studies discussed thus far focused on psychotic and mood disorders, Guggenmos et al. (2017) was the first to calculate the BAG in alcohol dependence. A multi-linear ridge regression trained on the GM of 97 HCs was used, and a positive BAG of +4 years (SEM = 0.7) was found in the alcohol dependence group (*n*=119), compared to HCs (*n*=97) whose brain age estimation was the same as their mean chronological age (43.7 years). A dose response relationship was observed, with the highest BAG in the oldest

alcohol dependent participants (+11.7 years, SEM = 2.4), after regressing out gender, lifetime alcohol consumption, smoking status, and general health status. No advanced brain ageing was detectable in the youngest alcohol dependent participants (ages 20-29).

1.3.7. Additional BAG Analyses

The following section will discuss additional analyses conducted in the articles included in the systematic review. Here, the BAG is compared across disorders and the association between the BAG and clinical characteristics, such as age of onset and medication use, is explored.

1.3.7.1. Comparisons Across Disorders. The BAG is more pronounced in psychotic disorders (specifically schizophrenia) than mood disorders. The BAG was higher in schizophrenia than MDD when compared to HCs in Koutsouleris et al. (2014) and Koutsouleris et al. (2015). In addition, the BAG was significantly higher in the schizophrenia group when directly compared to borderline personality disorder ($p < .007$) and at-risk mental states ($p < .001$; Koutsouleris et al., 2014; Koutsouleris et al., 2015). When compared to the at-risk group MDD had a significantly higher BAG ($p < .005$). In addition, both Nenadic et al. (2017) and Shahab et al. (2019) observed a significantly higher BAG in their schizophrenia groups compared to bipolar disorder ($p = .0097$ and $p < .001$, respectively). Tonneson et al. (2019) observed no significant difference between their schizophrenia and bipolar disorder groups.

1.3.7.2. Medication Use and Other Clinical Characteristics. Van Gestel et al. (2019) found a positive association between antipsychotic treatment and brain age in their bipolar disorder patients. In addition, a negative association between lithium treatment and brain age was observed in these patients. This association remained significant when controlling for current and past exposure to antipsychotics and anticonvulsants. In addition, a greater BAG was observed in patients currently using antipsychotic medication compared to

those who were not. The BAG was the highest in the non-lithium patient group taking antipsychotics (+5.90, 95% CI = [3.73, 8.07]) and the lowest in the lithium only patient group (–0.67 [–2.65, 1.32]). These findings could suggest a neuroprotective factor specific to lithium treatment. However, Hajek et al. (2019) found no differences in brain age scores between participants with or without lifetime history of lithium treatment. Alternatively, it could be interpreted that patients taking antipsychotics have a bigger BAG due to psychotic symptomatology, as Schnack et al. (2016) found a significant positive association between the BAG and antipsychotic dosage. Although, the latter explanation seems unlikely, as Van Gestel et al. (2019) and Shahab et al. (2019) found that brain age was not associated with psychotic symptoms, and Nenadic et al. (2017) did not observe advanced ageing in a sub-group of psychotic bipolar disorder patients.

In contrast, Chung et al. (2018) found that diagnostic group differences of BAG were not explained by antipsychotic medication use in the younger and older age groups and Koutsouleris et al. (2014) found no association between the BAG and psychotropic medications. While Han et al. (2020) found both medication-free patients and patients using medication at the time of the scan to have significantly higher BAGs than controls, these two sub-groups did not have a significantly different brain age from one another. Kolenic et al. (2018) observed greater brain age scores in both previously medication naïve FEP patients and previously medicated FEP patients compared to controls. Brain age scores were not associated with medication dosage at the time of scanning or with cumulative exposure to antipsychotics in this study. Moreover, brain age scores were not associated with treatment with antidepressants, anticonvulsants, or first-generation or second-generation antipsychotics, and were not associated with the number of different types of medications used. Shahab et al. (2019) did not assess these associations, but rather assessed whether current medication exposure was associated with brain structure, finding no significant associations.

Kaufmann et al. (2019) and Schnack et al. (2016) observed a negative association between BAG and global assessment of functioning scale in schizophrenia patients. In addition, Schnack et al. (2016) found that brain acceleration rate in schizophrenia patients was positively associated with total Positive and Negative Symptom Scale (PANSS) score at follow-up. Koutsouleris et al. (2014) observed a significant positive association between the BAG and total PANSS score and general PANSS score in HCs. In addition, a significant positive association was observed between BAG and total, positive and general PANSS score in patients with schizophrenia who had an illness duration of less than 1-year, were treated with antipsychotic agents for less than 1 year and had not received inpatient treatment.

Kambeitz-Ilankovic et al. (2019) assessed the predictive ability of two of 14 neurocognitive tests. The premorbid IQ model was unable to detect a discrepancy in the BAG between the CHR and HC groups and the Digit Symbol Substitution model failed to predict age in HCs. Hajek et al. (2019) and Kolenic et al. (2018) found that brain age scores were not associated with duration of untreated psychosis or duration of illness. Finally, Han et al. (2020) found that compared to controls significant BAGs were observed in remitted patients, patients with current MDD, patients with early age of onset, with mid-adult onset and with late-adult onset of MDD, and with both first-episode patients and recurrent patients. Of note, no significant difference between the subgroups of MDD patients were observed in post hoc analyses.

1.3.8. Synthesis of Study Results

In summary, whereas all psychiatric conditions under investigation showed evidence of advanced brain age in at least one published study, there were notable differences between the disorder groups. The BAG ranged between -1.25 (Nenadic et al., 2017) and +7.8 years (Shahab et al., 2019) across disorders in this systematic review. A BAG of greater than +3 years (a value which lies approximately half-way between the lowest BAG and the highest

BAG) was observed in mood disorders in only three studies out of seven (Koutsouleris et al., 2014; Koutsouleris et al., 2015; Van Gestel et al., 2019). In psychotic disorders, a BAG of greater than +3 years was observed in the main analyses of six studies (Kambeitz-Ilankovic et al., 2018; Kaufmann et al., 2019; Koutsouleris et al., 2014; Koutsouleris et al., 2015; Schnack et al., 2016; Shahab et al., 2019), and in the exploratory analyses of two (Kolenic et al., 2018; Nenadic et al., 2017). Alcohol dependence had a BAG greater than +3 years (Guggenmos et al., 2017), however this needs to be confirmed in future research. The publication dates varied from 2014-2019 and sample sizes from *n*=23 (Koutsouleris et al., 2015) to *n*=4314 (Han et al., 2020). MAE ranged from 2.21 years (de Nooij et al., 2019) to 6.84 years (Han et al., 2020) in mood disorders, 1.41 years (Chung et al., 2018) to 7.98 years (Tonneson et al., 2019) in psychotic disorders, and was 6.9 years in alcohol dependence (Guggenmos et al., 2017).

When comparing study design, Schnack et al. (2016) and de Nooij et al. (2019) stand out with their use of a longitudinal study design, which allowed them to capture the relationship between brain age and chronological age more reliably by measuring change over time. Schnack et al. (2016) found a decelerating ageing effect in their schizophrenia sample, particularly during the first five years after disease onset. The BAG had increased by approximately two years after the first year of disease onset and had slowed down to almost 0 after the first five years of follow-up. This coincided with ages 20-25, which is considered the end of a period of neuro-maturational processes (Cannon, 1997), suggesting an advanced neuromaturation in schizophrenia.

Methodologically speaking, only one paper calculated the BAG based on WM only (Van Gestel et al., 2019) with a few training models on both GM and WM (Hajek et al., 2019; Kaufmann et al., 2019; Kolenic et al., 2018; Shahab et al., 2019; Tonneson et al., 2019; Van Gestel et al., 2019). Model training sample sizes ranged from 36 (Kambeit-Ilankovic et

al., 2019) to 35,474 (Kaufmann et al., 2019). RVR was the most common brain age algorithm, specifically a model developed by Franke et al. (2010) which was used by Besteher et al. (2019), Hajek et al. (2019), Kolenic et al. (2018), Nenadic et al. (2017) and Van Gestel et al. (2019). RVR is a Bayesian alternative to SVR and unlike SVR it does not need to undergo parameter optimization training (Hajek et al., 2019). RVR is robust to differing scanner strengths, and it uses 'the most-typical' cases compared to SVR which uses 'non-typical' cases (Hajek et al., 2019; Kolenic et al., 2018). This may contribute to explaining why RVR is more stable and accurate than SVR in capturing whole-brain age related changes, and avoiding over-fitting (Hajek et al., 2019).

Ridge regression was the 2nd most common method (Chung et al., 2018; Guggenmos et al., 2017; Han et al., 2020), followed by v-SVR (Koutsouleris et al., 2014; Koutsouleris et al., 2015; Schnack et al., 2016). SVR can learn multivariate prediction rules from one sample and generalize this to a different sample allowing for unbiased age prediction, using cross-validation (Filzmoser et al., 2009). However, Schnack et al. (2016) reported that their model was insensitive to non-linear changes of GM with age. Shahab et al. (2019) used a random forest algorithm and Kaufmann et al. (2019) and Tonneson et al. (2019) used a xgboost model, trained separately on males and females. Of note, xgboost models have shown superiority in machine learning competitions and are known for their resource efficiency (Chen & Guestrin, 2016). Finally, all papers used T1-weighted structural MRI data, except for Tonneson et al. (2019; DTI data) and Shahab et al. (2019; structural MRI, DTI and cognitive performance scores).

The most common limitations include small psychiatric sample sizes (<50; Besteher et al., 2019; de Nooij et al., 2019; Hajek et al, 2019; Kambeitz-Ilankovic et al., 2019; Koutsouleris et al., 2015; Nenadic et al., 2017; Shahab et al., 2019; Van Gestel et al., 2019) and small HC training sample sizes (<200; de Nooij et al., 2019; Guggenmos et al., 2017;

Kambeitz-Ilankovic et al., 2019; Shahab et al., 2019). Kaufmann et al. (2019) tested the impact of sample size on model performance and found that as the sample size increased, so did the model performance (measured by the correlation between predicted brain age and chronological age). Other common study limitations include using a cross-sectional design precluding investigation of causal relationships (Kolenic et al., 2018; Koutsouleris et al., 2014; Van Gestel et al., 2019), conflation of effects of medication and disorder-specific effects on brain ageing (de Nooij et al., 2019; Nenadic et al., 2017; Schnack et al., 2016), limited information on clinical characteristics (Han et al., 2020), and using the same controls in analyses as used in model training (de Nooij et al., 2019; Guggenmos et al., 2017; Kambeitz-Ilankovic et al., 2019). This last limitation could cause inferential statistics to lose meaning as the test sample is not sufficiently independent from the selection sample (Kriegeskorte et al., 2009).

It is worth noting that brain age models are sensitive to the effects of disease. Therefore, the outcome may not capture brain age alone, but may also reflect additional factors, such as structural alterations as a result of disease itself (Chung et al., 2018; Kolenic et al., 2018; Koutsouleris et al., 2014; Nenadic et al., 2017; Schnack et al., 2016; Van Gestel et al., 2019). Moreover, the use of a single brain age value does not provide information on spatial specificity and disorder specific patterns (Kaufmann et al., 2019). Finally, the studies included in this review do not investigate the BAG alongside other measures of biological ageing, such as telomere length and epigenetic age. This is particularly important as different tissues, organs and systems in the human body may undergo different ageing trajectories (Cole, Marioni, et al., 2018). Therefore, by exploring and integrating other measures of ageing biomarkers, research will be able to gain a more comprehensive picture of advanced ageing in psychiatric disorders (Besteher et al., 2019).

1.3.9. Conclusions and Recommendations

A positive BAG is apparent across psychotic and mood disorders with the largest positive BAG observed in schizophrenia. Moreover, there was an attenuated effect seen in the initial stages of disease in both psychotic and mood disorders, as the BAG was not significant in participants who had recently converted to a disorder diagnosis. Future longitudinal research will be useful in identifying at what stage of the lifespan advanced ageing takes place. While a variety of methods were used to model the BAG, RVR was used the most frequently. This is likely due to its stability, accuracy and ability to avoid over-fitting. One of the most apparent gaps in the literature is the lack of papers investigating the BAG in anxiety disorders. It is imperative that future research explore brain ageing in anxiety disorders due to their high prevalence rates and debilitating nature (Kessler et al., 2005). While most papers explore the association between the BAG and clinical factors, very few consider future clinical applications of the BAG. Future studies should consider how this method could be used to track disease progression and identify risk of disease relapse. Overall, this is an important field of research that is only beginning to reveal the complex process of brain ageing in the context of psychiatric disease.

1.4. Research Questions and Hypotheses

Based on the findings discussed in the preceding systematic review, there is a gap in the literature regarding the BAG in anxiety disorders diagnosed in adults. This, in combination with evidence of advanced biological ageing in SAD, leads to the hypothesis that there will be a positive BAG in adult SAD patients as compared to HCs. To the best of our knowledge, the deviation between an individual's chronological age and their estimated brain age has not yet been investigated in adults diagnosed with SAD.

The following research questions will be investigated:

1. Is there a positive BAG present in patients diagnosed with SAD, as compared to HCs?

Hypothesis 1a: A statistically significant positive BAG will be found in the SAD group.

Hypothesis 1b: The BAG will be significantly larger in the SAD group compared to the HC group.

2. Does the BAG differ in SAD patients with comorbid psychiatric disorders, compared to patients without comorbid psychiatric disorders, and HCs?

Hypothesis 2: Patients with comorbid psychiatric disorders will have a larger positive BAG as compared to those without psychiatric comorbidities, and HCs.

3. Is the BAG larger in SAD patients with a higher symptom severity? Is the BAG associated with medication use?

Hypothesis 3: There will be a positive association between size of the BAG and higher self-reported symptom severity in SAD patients, as determined by LSAS score. The outcome will not be associated with medication use.

2. Methods

2.1. Study Design

Previously collected structural T1-weighted 3T MRI scans from 174 patients diagnosed with SAD, and 213 matched HCs, including both males and females, were used in this dissertation. The data were previously used in a VBM mega-analysis project (Bas-Hoogendam et al., 2017).

2.2. Participant Recruitment

The previously collected MRI scans were acquired at 10 research centres located in Africa, Europe and North America (Bas-Hoogendam et al., 2017). Participant recruitment methods differed at each site, as outlined below.

At the first site, participants were recruited at the University of Jena and University of Münster in Germany (Boehme, Mohr, et al., 2014; Boehme, Ritter et al., 2014; Boehme et al., 2015). For the Netherlands Study of Depression and Anxiety (NESDA), patients were recruited from outpatient clinics and regional mental health care facilities at Leiden University Medical Centre, VU Medical Centre in Amsterdam and University Medical Centre Groningen (Pannekoek et al., 2015; Penninx et al., 2008; van Tol et al., 2010). The community sample was recruited from two previously established cohorts (Bijl et al., 1998; Landman-Peeters et al., 2005). Participants were also recruited from Leiden University Medical Centre for the Social Anxiety Study (Cremers et al., 2014; Cremers et al., 2015) and the University of Illinois recruited patients from clinics (Klumpp et al., 2015). Finally, participants were recruited from the University of Cape Town (UCT) in South Africa (Geiger et al., 2016; Hattingh et al., 2013; Howells et al., 2015; Syal et al., 2012) and Umeå University in Sweden (Månsson et al., 2013). Permission for use of the participant data was granted by the study primary investigators. For information on the SAD and HC inclusion and exclusion criteria per study site, please refer to supplementary Table 1.

2.3. Measures

2.3.1. Clinical Measures

To obtain a primary diagnosis of SAD, the Structured Clinical Interview for DSM-IV Axis I and II disorders (SCID I and SCID II) (First et al., 1998) was used by the majority of sites, including University of Jena and University of Münster (Boehme, Mohr et al., 2014; Boehme, Ritter et al., 2014; Boehme et al., 2015), UCT (Geiger et al., 2016; Hattingh et al., 2013; Howells et al., 2015; Syal et al., 2012), Umeå University (Månsson et al., 2013) and University of Illinois (Klumpp et al., 2015). SCID has an inter-rater reliability ranging from adequate to excellent (Lobbestael et al., 2011). The Composite International Diagnostic Interview, lifetime version 2.1 (CIDI; Robins et al., 1988) was used in the NESDA study for SAD diagnosis (Pannekoek et al., 2015; Penninx et al., 2008; van Tol et al., 2010). The CIDI, developed by the World Health Organisation, is a diagnostic tool which takes approximately 2 hours to administer and has a high test-retest reliability (Kessler & Üstün, 2004). Finally, The Mini International Neuropsychiatric Interview, with high inter-rater and test-retest reliability, was used to obtain a primary diagnosis of SAD (Lecrubier et al., 1997) at Leiden University Medical Centre (Cremers et al., 2014; Cremers et al., 2015).

The Liebowitz Social Anxiety Scale (LSAS; Heimberg et al., 1999), a questionnaire with strong evidence in support of internal consistency, discriminant validity, convergent validity, and treatment sensitivity, was used to determine severity of SAD in most sites, including University of Jena and University of Münster (Boehme, Mohr et al., 2014; Boehme, Ritter et al., 2014; Boehme et al., 2015), Leiden University (Cremers et al., 2014; Cremers et al., 2015) and Umeå University (Månsson et al., 2013). LSAS was also used by University of Jena to aid in the diagnosis of SAD (Boehme, Mohr, et al., 2014; Boehme, Ritter et al., 2014; Boehme et al., 2015). Depression severity was measured with the Beck Depression Inventory (BDI) at all sites (Beck et al., 1998). The BDI is a reliable interview,

made up of 21 items of symptoms and attitudes formed from clinical observations of depressed patients (Beck et al., 1998).

2.3.2. Scan Parameters

The scan parameters at each study site are described in the below table.

Table 4.
Scan Parameters per Study Site

Country	Study site	Scanner	Voxels	Dimensions	Number of channels
Germany	University of Jena[1], University of Münster[1]	Siemens/TrioTim 3T	192 × 256 × 256	1 × 1 × 1 mm	Not available
The Netherlands	VU Medical Centre Amsterdam – NESDA study[2]	Philips 3T	170 × 256 × 256	1 × 1 × 1 mm	SENSE-6 channel head coil
	University of Groningen – NESDA study[2]	Philips 3T	170 × 256 × 256	1 × 1 × 1 mm	SENSE-8 channel head coil
	Leiden University Medical Centre – NESDA study[2]	Philips 3T	170 × 256 × 256	1 × 1 × 1 mm	SENSE-8 channel head coil
	Leiden University Medical Centre - Social Anxiety study[3]	Philips 3T	256 × 256 × 140	0.875 × 0875 ×1.2 mm	8-channel SENSE head coil
South Africa	University of Cape Town[4]	Siemens Magnetom Allegra 3T	128 × 256 × 256	1.33 × 1 × 1 mm	4-channel head coil
	Stellenbosch University[4]	Siemens Magnetom Allegra 3T	128 × 256 × 256	1.33 × 1 × 1 mm	4-channel head coil
Sweden	Umeå University[5]	General Electric 3T	512 × 512 × 176	0.48 × 0.48 × 1 mm	32-channel head coil
United States of America	University of Illinois[6]	GE Signa System 3T	256 × 256 × 182	0.86 × 0.86 × 1 mm	Standard radiofrequency coil
	University of Chicago[7]	GE Signa System 3T	256 x 256 x 120	0.94× 0.94× 1.5 mm	Standard radiofrequency coil
	University of Michigan[7]	GE Signa System 3T	256 x 256 x 124	1× 1× 1.2 mm	Standard radiofrequency coil

Note. Table adapted from Bas-Hoogendam et al. (2017). [1] Boehme et al., 2015, Boehme, Mohr et al., 2014, Boehme, Ritter et al.,2014. [2] Pannekoek et al., 2015; Penninx et al., 2008, 2013; van Tol et al., 2010. [3]Cremers et al., 2015, 2014. [4]Geiger et al., 2016; Hattingh et al., 2013; Howells et al., 2015; Syal et al., 2012. [5] Månsson et al., 2013. [6]Klumpp et al., 2015. [7]Phan et al., 2013.

2.4. Brain Age Models

2.4.1. Kaufmann Model

An established brain age model was used to predict brain age in the study participants (Kaufmann et al., 2019). The model was trained at the Norwegian Centre for Mental

Disorders Research on a large group of HCs (*n*=35,474; 18,990 females and 16,484 males) between the ages of 3 and 89 years (Kaufmann et al., 2019). It was trained separately on males and females using gradient tree boosting (Chen & Guestrin, 2016), to predict the age of the brain from FreeSurfer-derived cortical thickness, area and volume features from 180 ROIs, per hemisphere (Kaufmann et al., 2019). Features were extracted using a recent cortical parcellation scheme (please see Glasser et al., 2016). The classic set of cerebellar/subcortical and cortical summary statistics were also extracted (see Fischl et al., 2002) resulting in a total of 1118 structural brain imaging features. Out of sample testing of the training data was used to tune the model parameters, and for 1500 rounds the prediction error was evaluated to achieve the most suitable number of model training iterations. The remaining parameters utilised the standard settings for linear xgboost tree models. The final models were then trained on the full set of training data using the adjusted *nrounds* parameter (Kaufmann et al., 2019). The correlations between predicted brain age and chronological age within the training sample were high (female model: r =0.93, male model: r = 0.94). The correlations between predicted brain age and chronological age for the test datasets were also high: schizophrenia r = 0.72, prodromal schizophrenia or at-risk mental state r = 0.61, mixed diagnosis on the psychosis spectrum, r = 0.71, bipolar spectrum disorders r = 0.73 and MDD r = 0.77.

Three scripts were run on the dissertation participants to calculate the estimated BAG. The first script was run to extract features and the second script was run to create a feature subject matrix. Lastly, the brain age estimation script was run in R using the output from the previous two scripts, in addition to the following participant covariates: age and sex. Scan site could not be included as a covariate here, as the training and test samples used in the development of this model were collected at different research sites. However, Kaufmann et al. (2019) compared models trained within sites (individual HC groups for each diagnosis and

scanner) to models trained across sites (HC groups combined). Predicted brain age and the BAG from the within-site models were significantly correlated with the respective metrics obtained from across-site models, after accounting for age, age^2, sex and Euler number, with coefficients of comparable magnitude as observed in the present study. This confirms that the brain age model can be applied to independent sites without explicit modelling of site effects. In addition, scan site was included as a random factor in the general linear models (GLM) in this dissertation. This script generated the estimated brain age of the participants.

This brain age model was chosen for this project as it was expected to fit the data. Both the sample which the model was trained on and the participants used in this study are international samples, including both males and females from multiple scanner sites. Although the age range of the samples differ, the training dataset provides adequate coverage of the 18-60 age range, as the training dataset ranged from age 3-89 years. Furthermore, the model was chosen due to prior evidence of superior performance and resource efficiency of the xgboost package in R (Chen & Guestrin, 2016). Of note, due to model bias concerning correlations between chronological age and BAG, it is recommended to include age as a covariate in analyses which use the BAG as an outcome measure (Cole et al., 2015). This is the approach used in the present dissertation. However, Kaufmann model fit with the data used in this dissertation was sub-optimal. MAE was high, as the comparison group (HCs) has a MAE that is substantially greater than 0: whole sample (*n*=238) = 11.93 (6.93), SAD patients only = 13.33 (7.03) and HCs only = 10.91 (6.7). The Pearson's correlations and R^2 between chronological age and predicted brain age were moderate when all research sites were grouped together whole sample, r = 0.576 (R^2 = 0.332), patients only, r = 0.501 (R^2 = 0.251) and HCs, r = 0.627 (R^2 = 0.393), and were less favourable for certain research sites (NESDA Leiden r = 0.269; CUBIC r = 0.219). Therefore, it was decided that a different brain age model with better model fit statistics should be used in addition to the Kaufmann model.

2.4.2. Cole Model

A second model chosen for the present dissertation is James Cole's BrainageR (Cole et al., 2017; Cole et al., 2015; Cole & Ritchie et al., 2018; Karatzoglou et al., 2004). The Cole model predicts brain age from raw T1-weighted MRI scans with a Gaussian Processes regression in R, using the kernlab package. The software uses SPM12 for segmentation and normalisation, and *slicesdir* was used to generate PNG images and an index.html file for quality checking purposes. Lastly, the normalised images are uploaded into R. The Cole model was trained on 3377 healthy individuals ages 18-92 from seven publicly available datasets: Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing (https://aibl.csiro.au/), Dallas Lifespan Brain Study (http://fcon_1000.projects.nitrc.org/indi/retro/dlbs.html), Brain Genome Superstruct Project (https://dataverse.harvard.edu/dataverse/GSP), IXI (https://brain-development.org/ixi-dataset/), Nathan Kline Institute Rocklands Sample Enhanced (http://fcon_1000.projects.nitrc.org/indi/enhanced/), Open Access Series of Imaging Studies-1 (https://www.oasis-brains.org/) and Southwest University Adult Lifespan Dataset (http://fcon_1000.projects.nitrc.org/indi/retro/sald.html).

The Cole model was then tested on a sub-sample of the 3377 participants (*n*=857), with random assignment to both test and training data. Of note, the model does not correct brain age predictions for statistical dependency on chronological age. Model fit statistics in the held-out test sample were as follows: r = 0.973, MAE = 3.93 and R^2 = 0.946. Correlation between chronological age and BAG was r = - 0.012. In a completely independent sample (CamCAN, *n*=611, age 18-90 years), model fit statistics were as follows: r = 0.947, MAE = 4.90 years. Correlation between age and BAG was - 0.379 in this dataset.

2.5. Image Processing and Quality Control Procedure

2.5.1. Kaufmann Model

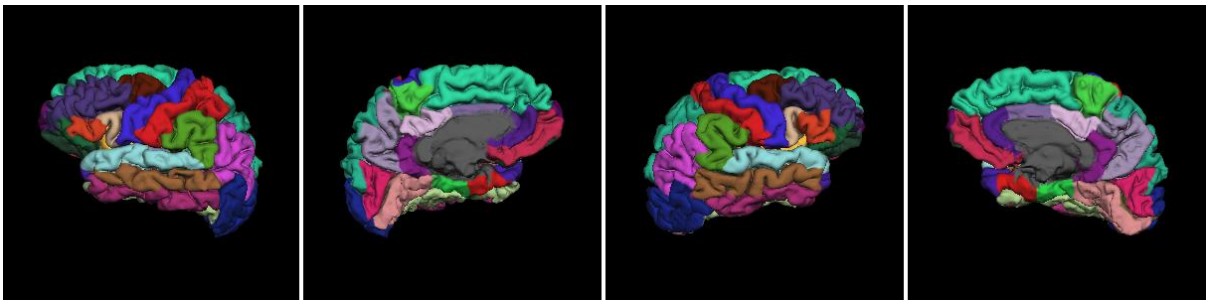
For the original xgboost model, the raw T1 scans of the SAD sample were processed using FreeSurfer version 5.3 by performing subcortical segmentation and surface-based morphometry. Surface area and thickness values were extracted using FreeSurfer, generating 4 external images per participant (2 lateral and 2 mid-sagittal images), and 8 internal images per participant (please see Figure 2). Participants were excluded from the study when FreeSurfer failed to output labels, or if there was a general segmentation failure. Segmentation failed for all scans of two study sites (n = 124; University of Chicago and University of Michigan) as the correct spatial orientation of the converted scan images could not be confirmed. The raw data was requested from the relevant investigators, however they did not share these with us for inclusion in the present dissertation. The segmentations were then assessed independently by myself and a reviewer (AA). AA has experience performing QC, and I was trained by my co-supervisor (JI) in the use of a visual quality control (QC) procedure according to established Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) protocols (http://enigma.ini.usc.edu/protocols/imaging-protocols/; Stein et al., 2012). The protocol consisted of checking for a variety of segmentation errors, including general segmentation failure due to processing errors or severe global misclassifications, imaging artefacts due to motion, misclassifications of brain regions, and over/underestimations of brain regions. In addition, summary statistics and histograms were generated, allowing for the identification of outliers.

QC was focused on cortical regions as most features used in the Kaufmann model are cortical (Kaufmann et al., 2019). Moreover, Kaufmann et al. (2019) used FreeSurfer's Euler number as the primary method for QC in their study. The Euler number measures the topological complexity of the cortical surface of MRI scans (Rosen et al., 2018). Higher negative values of the Euler number indicate poorer scan quality. The formula used to convert the number of surface holes into a Euler number is as follow: 2-2n, where n is

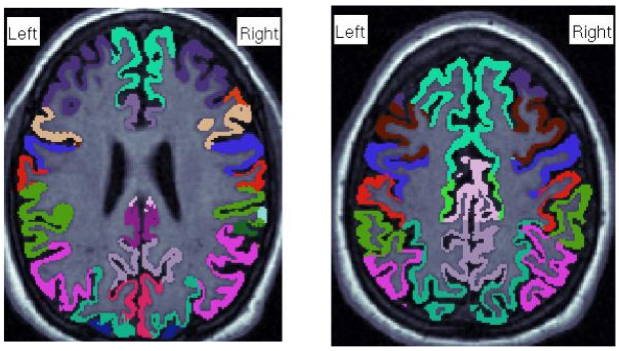
replaced by the number of surface holes. This is done for the left and right hemisphere, after which the left and right hemisphere Euler numbers are averaged to create one Euler number per participant (Rosen et al., 2018).

Figure 2.
FreeSurfer External and Internal Views of Segmentations

A



B



Note. Figure adapted from Ching et al. (2017). http://enigma.usc.edu/ ; Stein et al. (2012). Image A shows external segmentations. Image B shows internal segmentations.

The independent reviewer, my one supervisor and I added additional specifications to the existing QC procedure to ensure consistent interpretation for final QC. For inter-rater reliability please see section 2.2 of the supplementary materials. Participants were excluded if they received seven or more QC 'flags' by both raters or if they were an outlier for seven or more brain regions, as this resulted in a loss of at least 10% of participant data across both hemispheres. Twenty-three participants were excluded based on number of flags and outlier regions, resulting in a total of 240 participants, 101 of which had a diagnosis of SAD and 139

of which were HCs. For more information on FreeSurfer QC exclusion, please see the supplementary materials (supplementary Table 2.1).

Lastly, whole-brain Euler number was calculated for each participant. Mean Euler number was then calculated per research site, and two participants (one SAD patient from CUBIC and one HC participant from the University of Münster) were excluded due to being three standard deviations below the mean Euler number. The final sample to be used in the Kaufmann analyses consisted of a total of 238 participants, 100 of which had a diagnosis of SAD and 138 of which were HCs. Participant demographic and clinical information can be seen in Table 5 below.

Table 5.
Kaufmann Model – Participant Demographic and Clinical Information from 8 Research Sites

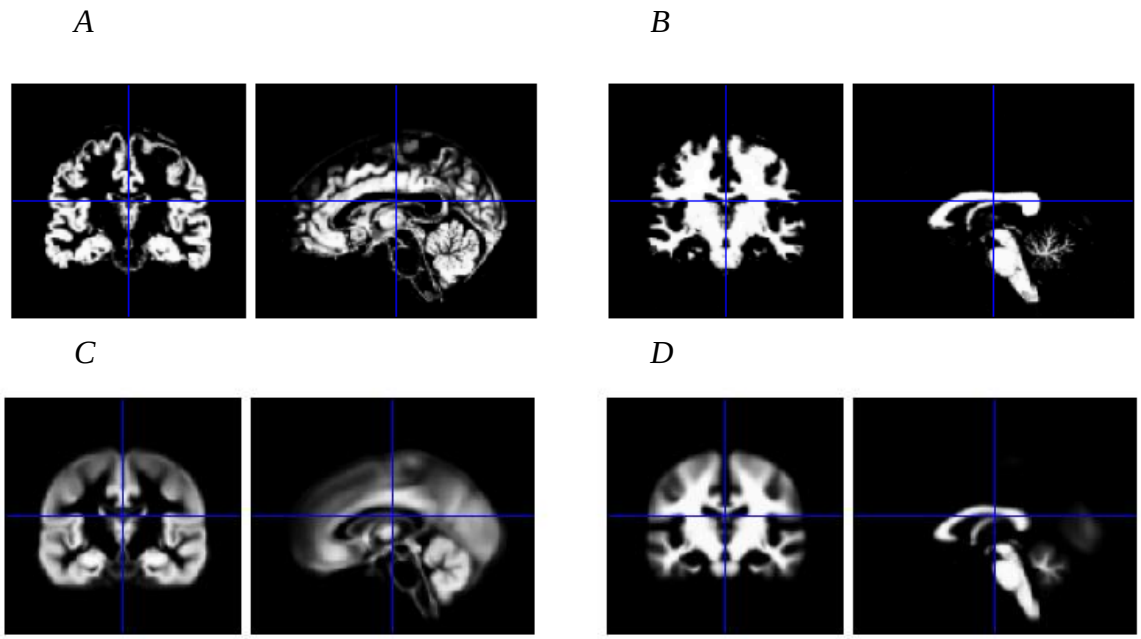
	All participants	**SAD (*n*=100)**	**HC (*n*=138)**	**[7]Research site differences (one-way ANOVAs)**
Female n (%)	140 (58.82%)	61 (61.00%)	79 (57.20%)	$p < .001$
Mean (sd) age	33.38 (10.31)	31.71 (9.69)	34.59 (10.61)	$p < .001$
[1]Education n (%)				$p = .031$
Low (5-8y)	3 (1.56%)	1 (1.25%)	2 (1.79%)	
Intermediate (9-14y)	59 (30.73%)	26 (32.5%)	33 (29.46%)	
High (15y+)	130 (67.71%)	53 (66.25%)	77 (68.75%%)	
Right-handed n (%)	232 (97.5%)	99 (99%)	133 (96.4%)	$p = .014$
[2]Median (IQR) AOO	14 (10)	14 (10)	N/A	SAD only: $p = .214$
[3]Comorbidity (n)	19	19	N/A	SAD only: $p = .036$
MDD (n)	9	9		
Anxiety (n)	12	12		
[4]Medication use n (%)	22 (23.16%)	22 (23.16%)	N/A	SAD only: $p < .001$
[5]Mean (sd) LSAS	48.84 (34.12)	78.29 (18.85)	18.61 (13.48)	SAD: $p = .025$ HC: $p = .216$
[6]Mean (sd) BDI	10.18 (9.62)	16.50 (9.68)	4.00 (3.82)	SAD: $p = .247$ HC: $p = .010$

Note. AOO, age of onset; BDI, Beck Depression Inventory; LSAS, Liebowitz Social Anxiety Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. [1]Education is available for 192 participants (SAD n =80, HC n = 112). [2]AOO is available for 58 participants, [3]MDD and Anxiety do not add up to total comorbidity due to participants having both comorbidities. [4]Medication use is available for 95 participants. [5]LSAS is available for 154 participants (SAD n = 78, HC n = 76). [6]BDI is available for 93 participants (SAD n = 46, HC n = 47). [7]In the whole sample, Levene's test was violated for gender, education and handedness. In SAD patients, Levene's test was violated for comorbidity and medication use. ANOVA results may be overestimated due to the violations of inhomogeneity of variances assumption.

2.5.2. Cole Model

QC was redone for the Cole model, as this model uses statistical parametric mapping (SPM) to segment GM, WM, and cerebrospinal fluid (CSF) from 3D scans. SPM generates GM maps, WM maps and CSF maps in the participants' original native space and in template space (please see Figure 3). The maps in the template space have been normalised to MNI space (Cole et al., 2017). Like the previous QC protocol, the original sample of 387 participants was checked for misclassification, thickness and folding errors (Ashburner, 2010). An additional QC step, misregistration, was checked in the normalised scans to ensure there were no issues with scan alignment to MNI space.

Figure 3.
SPM GM, WM and CSF Segmentations in Native Space and MNI Space



Note. Figure adapted from Ashburner et al. (2010). GM, grey matter; MNI, Montreal Neurological Institute; WM, white matter; SPM, statistical parametric mapping. Images in the top row are in native space. Images in the bottom row are in MNI space. Image A and C show GM maps. Image B and D show WM maps.

The new QC largely corresponded to the FreeSurfer QC, where many of the participants who were failed previously were failed once again. However, there were instances in which participants who were failed previously have now been retained (*n*=17) and participants who were passed previously were excluded from the final sample (*n*=10),

confirming the necessity of redoing QC for the Cole model. Participants who were failed and who were classified as being on the border between pass and fail were checked by my supervisor and a final decision was made. For more information on SPM QC exclusion please see the supplementary materials (supplementary Table 2.2). The final sample to be used in the Cole analyses consisted of a total of 352 participants (SAD *n*=155, HC *n*=197). Demographic and clinical information can be seen in Table 6 below.

Table 6.
Cole Model – Participant Demographic and Clinical Information from 10 Research Sites

	All participants	**SAD (*n*=155)**	**HC (*n*=197)**	**[7]Research site differences (one-way ANOVAs)**
Female n (%)	200 (56.82%)	91 (58.70%)	109 (55.30%)	$p = .001$
Mean (sd) age	31.26 (10.06)	30.40 (9.40)	31.93 (10.52)	$p < .001$
[1]Education n (%)				$p = .049$
Low (5-8y)	7 (2.29%)	1 (0.74%)	6 (3.51%)	
Intermediate (9-14y)	98 (32.03%)	46 (34.07%)	52 (30.41%)	
High (15y+)	201 (65.69%)	88 (65.19%)	113 (66.08%)	
Right-handed n (%)	345 (98%)	154 (99.40%)	191 (97%)	$p = .017$
[2]Median (IQR) AOO	14 (10)	14 (10)	N/A	SAD only: $p = .385$
[3]Comorbidity (n)	32	32	N/A	SAD only: $p = .004$
MDD (n)	10	10		
Anxiety (n)	24	24		
[4]Medication use n (%)	22 (14.57%)	22 (14.57%)	N/A	SAD only: $p < .001$
[5]Mean (sd) LSAS	45.7 (35.11)	77.28	13.88	SAD: $p = .025$ HC: $p < .001$
[6]Mean (sd) BDI	7.94 (8.78)	13.79 (8.98)	2.36 (3.21)	SAD: $p = .011$ HC: $p < .001$

Note. AOO, age of onset; BDI, Beck Depression Inventory; LSAS, Liebowitz Social Anxiety Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. [1]Education is available for 306 participants (SAD n = 135, HC n = 171). [2]AOO is available for 59 participants, [3]MDD and Anxiety do not add up to total comorbidity due to participants having both comorbidities. [4]Medication use is available for 151 participants. [5]LSAS is available for 261 participants (SAD n = 131, HC n = 130). [6]BDI is available for 205 participants (SAD n = 100, HC n = 105). [7]In the whole sample, Levene's test was violated for gender, age, education and handedness. In SAD patients, Levene's test was violated for comorbidity and medication use. In HCs, Levene's test was violated for LSAS and BDI. ANOVA results may be overestimated due to the violations of inhomogeneity of variances assumption.

2.6. Statistical Analysis

The generalisability of the brain age models were determined by applying the trained brain age algorithms to the HC groups. Model fit statistics were calculated for both brain age

models using Pearson correlation coefficients and explained variance (R^2) between estimated brain age and chronological age in the HC group, the SAD group and the whole sample. To assess model performance, the MAE was calculated between estimated brain age and chronological age of participants in both groups and in the whole sample and was plotted for inspection of MAE distributions per site.

To calculate the BAG for each participant, chronological age was subtracted from model predicted brain age. The BAG was used as the outcome variable in subsequent statistical analyses. T-tests were run to determine whether the SAD group and HC group mean BAG was significantly different than 0. To examine the main effect of SAD diagnosis GLMs were run on the combined dataset, adjusting for possible confounding variables, including sex, age and mean centred age^2. Research site was included as a random factor in the GLMs. In addition, a GLM was run to determine if BAG differs in SAD patients with comorbid psychiatric disorders, SAD patients without comorbidities and HCs. Once again diagnosis was included in the model as a fixed factor, research site as a random factor, and sex, age and mean centred age^2 were included as covariates. Post hoc comparisons with Bonferroni correction for multiple comparisons were run to identify any BAG mean differences between the three groups. Lastly, a GLM was run in SAD patients who had information on symptom severity (LSAS) and medication use. These two variables were included in the model as additional covariates with diagnosis as a fixed factor and research site as a random factor. All analyses were run in both the Cole model dataset and the Kaufmann model dataset. Mega-analyses and descriptive statistics were performed in SPSS v25. Throughout analyses, statistical significance is defined according to convention, as $p < .05$.

2.7. Ethics Statement

Study approval was granted by the Human Research Ethics Committee of the Faculty of Health Sciences, UCT (reference: 303/2019). At each site, separate ethical approval for

data collection was obtained in accordance with the site-specific study protocol (Bas-Hoogendam et al., 2017). As such, this MSc project complies with the declaration of Helsinki (World Medical Association, 2013). Participant data has been fully de-identified and access to the project data has been restricted to the researchers of this study and the investigators based at Leiden University (Bas-Hoogendam et al., 2017).

3. Results

3.1. Model Fit

Table 7 shows the model fit statistics for the two brain age models presented in this dissertation, including Pearson correlations between brain age and chronological age, R^2 between brain age and chronological age, and the mean BAG and MAE. Based on these statistics, brain age estimates from the Cole model were closer to chronological age, as the MAE is closer to 0 (Cole: whole sample MAE = 4.78, HC MAE = 4.79 versus Kaufmann: whole sample MAE = 11.93, HC MAE = 10.91). When correlated with one another, the Cole model and the Kaufmann model demonstrated a moderate Pearson correlation of 0.511 ($p <$ .001) for estimated brain age, and there was no significant association for BAG (Pearson correlation = 0.120, $p = .066$).

Table 7.
Model Fit Statistics for Cole Model and Kaufmann Model in the Combined Sample and Separately in SAD and HC Participants

	Cole whole sample (*n*=352)	**Cole SAD (*n*=155)**	**Cole HC (*n*=197)**	**Kaufmann whole sample (*n*=238)**	**Kaufmann SAD (*n*=100)**	**Kaufmann HC (*n*=138)**
*Pearson correlation	0.828**	0.830**	0.826**	0.576**	0.501**	0.627**
R^2	0.686 (6.12)	0.688 (5.95)	0.683 (6.27)	0.332 (3.48)	0.251 (3.28)	0.393 (3.58)
Mean BA gap (SD)	-0.39 (6.20)	-0.58 (5.96)	-0.25 (6.39)	10.79 (8.60)	12.48 (8.46)	9.57 (8.52)
MAE (SD)	4.78 (3.96)	4.77 (3.60)	4.79 (4.23)	11.93 (6.93)	13.33 (7.03)	10.91 (6.95)

Note. Pearson correlations: **p < .001. BA, brain age; HC, healthy control; MAE, mean absolute error; SAD, social anxiety disorder; SD, standard deviation.
*Pearson correlations are between brain age and chronological age.

3.2. Main Analyses: The BAG in SAD Patients and HCs

Section 5 of the supplementary materials demonstrates that there were no violations of assumptions of normality, homogeneity of regression slopes and homogeneity of variance for the analyses in this dissertation.

3.2.1. Hypothesis 1a: A Statistically Significant Positive BAG will be Found in the SAD Group

The first part of the main analysis aimed to determine if there is a positive BAG in the SAD group and HC group.

3.2.1.1. Kaufmann Model. The findings from the Kaufmann model confirmed the current dissertation's hypothesis that the SAD group (n=100) had a positive BAG (M = +12.48, SD = 8.46) that was different from 0 (t(99) = 14.756, p < .001). However, the HC group (n=138) also demonstrated a significant positive BAG (M = +9.57, SD = 8.52, t(137) = 13.190, p < .001).

3.2.1.2. Cole Model. Contrary to expectations, a one-sample t-test showed that the SAD group (n=155) did not have a positive BAG (M = -0.58, SD = 5.96) that was significantly different from 0 (t(154) = -1.203, p = .231). Similarly, the HC group (n=197) did not have a BAG (M = -0.25, SD = 6.39) that was different from 0, t(196) = -0.539, p = .590.

3.2.2. Hypothesis 1b: The BAG will be Significantly Larger in the SAD Group Compared to the HC Group

The second part of the main analysis aimed to determine if there is a significantly larger BAG in the SAD group compared to the HC group.

3.2.2.1. Kaufmann Model. Table 8A demonstrates that, in line with expectations, diagnosis was significant (β = 0.943, p = .019, η_p^2 = 0.024). The SAD group had a significantly bigger BAG than the HC group of almost one year as seen in Figure 4 (mean difference = 0.943, SE = 0.40, p = .019). A secondary analysis was run, including a diagnosis-by-age interaction in the GLM, however there were no significant findings (β = 0.031, p =.432, η_p^2 = .003).

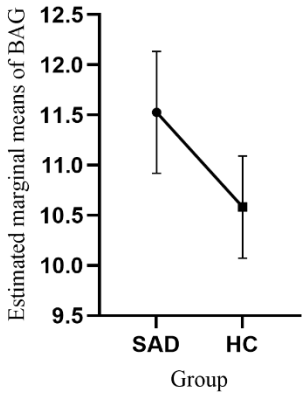
Table 8A.
GLM Investigating Between-Group Differences (SAD and HC) in BAG, Kaufmann Model

Variable	Type III SS	df factor	df error	F	p	η_p^2
Intercept	18436.104	1	65.511	1597.563	p < .001	0.961
Sex	141.886	1	226	16.525	p < .001	0.068
Age	9714.396	1	226	1131.391	p < .001	0.834
Mean centred age^2	182.988	1	226	21.312	p < .001	0.086
Diagnosis	48.071	1	226	5.599	.019	0.024
Site	387.248	7	226	6.443	p < .001	0.166

Note. Intercept SS error = 756.003, remaining SS errors = 1940.491. df, degrees of freedom; GLM, general linear model; SS, sum of squares error. BAG as the outcome variable, diagnosis as fixed factor, site as random factor and sex, age and mean centred age^2 as covariates.

Figure 4.

Estimated Marginal Means of BAG in SAD Patients and HCs, Kaufmann Model



Note. Error bars: 95% CI. HC, healthy control; SAD, social anxiety disorder. Covariates included in the model: sex, age and mean centred age^2.

3.2.2.2. Cole Model. Contrary to expectations, SAD patients did not show a larger BAG than HC participants (β = 0.513, p =.383, η_p^2 = .002). Although the SAD group demonstrated a higher adjusted mean BAG than the HC group (as seen in Figure 5), this difference was not significant (mean difference = 0.513, *SE* = 0.49, p = .383) as can be seen in Table 8B. In addition, a secondary analysis was run, including a diagnosis-by-age interaction in the GLM. A borderline significant diagnosis-by-age interaction was observed (β = -0.116, p =.050, η_p^2 = .011). Exploratory post hoc analyses were run to determine if any group differences in BAG between SAD patients and HCs could be detected in varying age groups as defined by age tertilesNo significant group differences were found between SAD patients and HCs in the youngest (participants ages 18-25, SAD *n*=56, HC *n*=73) and oldest (participants ages 36+, SAD *n*=41, HC *n*=70) age groups (youngest group: β = -1.528, p =.074, η_p^2 = .027. Oldest group: β = 1.985, p =.140, η_p^2 = .022). SAD patients had a significantly larger BAG compared to HCs in participants aged 26-35 years (SAD n=58, HC n=54, β = 2.11, p < .05, η_p^2 = .048, group mean difference = 2.11.

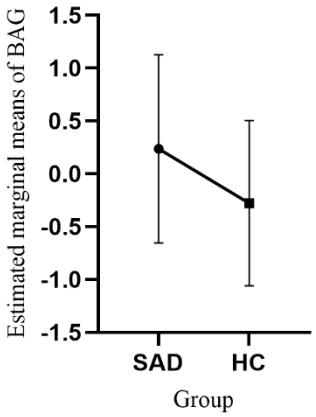
Table 8B.
GLM Investigating Between-Group Differences (SAD and HC) in BAG, Cole Model

Variable	Type III SS	df factor	df error	F	*p*	η_p^2
Intercept	531.052	1	43.646	11.100	.002	0.203
Sex	35.776	1	338	1.273	.260	0.004
Age	801.365	1	338	28.506	p < .001	0.078
Mean centred age^2	124.468	1	338	4.428	.036	0.013
Diagnosis	21.453	1	338	0.763	.383	0.002
Site	3454.553	9	338	13.654	p < .001	0.267

Note. Intercept SS error = 2088.062, remaining SS errors = 9501.846. df, degrees of freedom; GLM, general linear model; SS, sum of squares error. BAG as the outcome variable, diagnosis as fixed factor, site as random factor and sex, age and mean centred age^2 as covariates.

Figure 5.

Estimated Marginal Means of BAG in SAD Patients and HCs, Cole Model



Note. Error bars: 95% CI. HC, healthy control; SAD, social anxiety disorder. Covariates included in the model: sex, age and mean centred age^2.

The main analysis was rerun in the Cole database after removing the two sites (University of Chicago and University of Michigan) that had been excluded during QC in the Kaufmann database, to improve the comparability of the results of the two models. With the restricted sample, the BAG in SAD patients was still not significantly different from 0, and the BAG did not differ significantly between the SAD and HC groups. Please see section 7 of the supplementary materials for additional information on this analysis.

3.3. Sub-analysis 1: Psychiatric Comorbidities

Here, GLMs were run to determine whether there is a difference in BAG between SAD patients with comorbid psychiatric disorders, SAD patients without comorbidities and HCs.

3.3.1. Kaufmann Model

Comorbidity recoded was set as the fixed factor for this analysis (SAD patients with one or more comorbidities: *n*=19: MDD: *n*=9, other anxiety disorder: *n*=12, research site *n*=5, SAD patients without comorbidities: *n*=57, research site *n*=7, and HCs: *n*=138, research site *n*=8). Contrary to the hypothesis, no significant result was found for the SAD group with comorbidities (β = 0.258, *p* = .744, η_p^2 = 0.001). However, a significant result was found for the SAD group without comorbidities (β = 1.242, *p* = .013, η_p^2 = 0.031). SAD patients without comorbidities had a significantly higher BAG than HCs by more than one year (mean difference = 1.242, *SE* = 0.49, *p* = .038). Of note, the groups in this analysis (particularly the SAD with comorbidities group) are small, possibly limiting power to detect between-group differences.

3.3.2. Cole Model

The same general-linear model was run here (SAD patients with one or more comorbidities: *n*=32: MDD: *n*=10, other anxiety disorder: *n*=24, research site *n*=8), SAD patients without comorbidities: *n*=100, research site *n*=9, and HCs: *n*=197, research site *n*=10). In contravention of the hypothesis, no significant results were found – SAD without comorbidity: β = 0.549, *p* = .425, η_p^2 = 0.002; SAD with comorbidity: β = 1. 275, *p* = .236, η_p^2 = .004.

3.4. Sub-analysis 2: Symptom Severity and Medication Use

The final analysis sought to determine if there is a positive association between the BAG and SAD symptom severity, and if the BAG is associated with medication use.

3.4.1. Kaufmann Model

A general linear model was run in SAD patients who had information on symptom severity and medication use available (*n*=76, research site *n*=5) to assess the association between BAG and SAD symptom severity and medication use. Table 9A shows that no significant results were found (LSAS: β = 0.026, *p* = .212, η_p^2 = 0.023; medication use: β = -

1.407, p = .208, η_p^2 = 0.024). In addition, the model including BDI as a covariate (n=46, research site n=3) was once again not significantly associated with BAG (β = -0.015, p =.813, η_p^2 = 0.002).

Table 9A.
GLM Investigating the Association Between BAG and Symptom Severity and BAG and Medication use, Kaufmann Model

Variable	Type III SS	df factor	df error	F	p	η_p^2
Intercept	2386.061	1	67.737	242.152	p < .001	0.781
Age	2757.489	1	66	300.005	p <.001	0.820
Mean centred age^2	77.383	1	66	8.419	.005	0.113
LSAS	14.597	1	66	1.588	.212	0.023
Medication use	1.837	1	66	1.641	.208	0.024
Sex	15.270	1	66	1.661	.202	0.025
Site	113.066	4	66	3.075	.022	0.157

Note. Intercept SS error = 667.449, remaining SS errors = 606.636. df, degrees of freedom; GLM, general linear model; SS, sum of squares error. BAG as the outcome variable, sex as Fixed Factor, site as random factor and age and mean centred age^2, LSAS score and medication use as covariates.

Of note, all Kaufmann GLMs were run a second time using mean centred Euler number as a covariate. It was not included in the Cole model, as it is more appropriate to include the Euler number in analyses using cortical thickness data rather than voxel-based data as the Euler number measures the topological complexity of the cortical surface of MRI scans (Rosen et al., 2018). However, the Euler number did not make a difference to the results and was removed from all analyses. In addition, all Kaufmann GLMs were run with mean-centred age included as a covariate in place of age. Overall, there were no substantial changes to the results, other than to the intercept values.

3.4.2. Cole Model

The same general linear model was run in SAD patients who had information on symptom severity and medication use available (n=130, research site n=7). Like in the Kaufmann model analysis, Table 9B shows that no significant results were found for the above analysis (LSAS: β = -0.005, p = .855, η_p^2 = 0.000; medication use: β = -0.657, p = .688, η_p^2 = 0.001). A post hoc model including BDI total score as a covariate was run

(*n*=100, research site *n*=5), however it was not significantly associated with BAG (β = 0.027, p =.635) and was subsequently removed from the analysis.

Table 9B.
GLM Investigating the Association Between BAG and Symptom Severity and BAG and Medication use, Cole Model

Variable	**Type III SS**	**df factor**	**df error**	**F**	***p***	η_p^2
Intercept	53.686	1	88.415	2.133	.148	0.024
Age	171.587	1	118	8.137	.005	0.065
Mean centred age^2	26.699	1	118	1.408	.238	0.012
LSAS	0.710	1	118	0.034	.855	0.000
Medication use	3.414	1	118	0.162	.688	0.001
Sex	13.305	1	118	0.631	.429	0.005
Site	1047.451	6	118	8.279	p < .001	0.296

Note. Intercept SS error = 2225.409, remaining SS errors = 2488.228. df, degrees of freedom; GLM, general linear model; SS, sum of squares error. *BAG as the outcome variable, sex as fixed factor, site as random factor and age and mean centred age^2, LSAS score and medication use as covariates.*

Of note, all Cole GLMs were run with mean-centred age included as a covariate in place of age. Overall, there were no substantial changes to the results, other than to the intercept values.

4. Discussion

As revealed by the systematic review included in this dissertation, there is a gap in the literature regarding brain ageing in adults diagnosed with anxiety disorders. Based on this, the main aim of the present dissertation was to test the hypothesis that individuals with a diagnosis of SAD would demonstrate a positive BAG relative to HCs. Unexpectedly, the Cole model did not find a significantly larger positive BAG than 0 in SAD patients, while the Kaufmann model did. However, the Kaufmann model also found that the HC group BAG was significantly larger than 0, introducing concern about the validity of the model's predictions in our sample. Perhaps more telling was the lack of a significant difference in size of the BAG between SAD patients and HCs in the Cole model. In contrast, the Kaufmann model found that the SAD group had a significantly larger BAG than HCs, suggesting a possible advanced brain ageing in adults diagnosed with SAD.

Additionally, the present study aimed to determine whether the BAG would be larger in SAD patients with psychiatric comorbidities, higher SAD symptom severity, and medication use. A significantly larger BAG in SAD patients without psychiatric comorbidities, as compared to HCs, was found in the Kaufmann model. Again, no significant findings were observed in the Cole model. Of note, the Cole model demonstrated superior model fit statistics to the study data compared to the Kaufmann model.

4.1. The BAG: SAD and HCs

SAD patients had a mean BAG of +12.48 (8.46) years and HCs of +9.57 (8.52) years according to the Kaufmann model and of -0.58 (5.96) years and HCs of -0.25 (6.39) years according to the Cole model. However, when GLMs were run adjusting for covariates, a mean difference in BAG between SAD and HCs of +0.943 years (p = .019) was observed in the Kaufmann model and of +0.513 years (p = .383) in the Cole model. The confidence intervals for the group mean difference in BAG for the Kaufmann model (lower bound =

0.158, upper bound = 1.728) and the Cole model (lower bound = -0.642, upper bound = 1.668) were relatively similar, with a wider interval observed in the Cole model. Although the group difference was significant in the Kaufmann model but not the Cole model, the substantial overlap in the confidence intervals suggests that both models detected a similar mean difference in BAG between SAD patients and HCs. This suggests that if there is a positive BAG in adults diagnosed with SAD compared to healthy individuals, it is likely to be small as indicated by the confidence intervals provided by the two models.

The small positive BAG observed in the Kaufmann model could be interpreted in multiple ways. The finding suggests that the SAD group may have an increased risk of developing brain-related diseases due to an accelerated brain ageing process, as older individuals may be at an increased risk of developing health-related problems and worsening of disease symptoms (Cole & Franke, 2017; Cole, Marioni, et al., 2018). This interpretation is further supported by the diagnosis-by-age interaction that was observed in the Cole model. This finding suggests that age-related changes to the brain may be more apparent as an individual grows older. Indeed, there was a significantly larger BAG in SAD patients compared to HCs in an analysis including participants aged 26-35 (+2.11 years), suggesting these age-related changes may start after the brain has fully matured (Johnson et al., 2009; Sowell et al., 2003). It is important to note that while it is possible that what we observe is an acceleration of ageing, this cannot be statistically tested in this dissertation due to the cross-sectional design.

There are multiple possibilities that could explain the mechanistic relationships between SAD and brain ageing. The disorder itself may have had a role in advancing ageing-related changes to brain structure (Cole et al., 2019; Cole, Marioni, et al., 2018). One study found a relationship between brain ageing and inflammation (Le et al., 2018). Ibuprofen, an anti-inflammatory medication, was found to reduce the BAG by more than one year in

healthy participants, compared to a placebo group. In addition, there is evidence that both anxiety disorders and biological ageing have a relationship with oxidative stress. Oxidative stress contributes to premature cell death, which is associated with degenerative diseases and psychiatric disturbances (Tsaluchidu et al., 2008). Future research should aim to uncover the molecular and cellular processes driving the structural alterations to the brain that are captured by brain ageing models (Cole, Marioni, et al., 2018). This will help us to understand whether the alterations observed arise directly from biological ageing of the brain, or if they are associated with epigenetic factors involved in brain development and plasticity, such as inflammation (Han et al., 2020).

To the best of my knowledge, there is only one other study which investigates the BAG in anxiety disorders (Niu et al., 2019). The study included adolescents with a variety of psychiatric disorders, such as specific phobia, PTSD, and social phobia (*n*=142). This paper was not included in the systematic review due to the age of the social phobia group (*M* age = 14.46). A brain age model trained using SVR detected a positive BAG of similar value to that observed in the present study compared to controls (+0.55 years). However, after correcting for systematic gender and age biases commonly seen in brain age models the BAG was no longer significant. Limitations to the Niu et al. (2019) study include a small control group (*n*=60) and a lack of an independent dataset from the test sample of HCs for model training.

Compared to other psychiatric disorders, the BAG observed in the current study is relatively small. As demonstrated in the systematic review, the largest BAGs are apparent in psychotic disorders, particularly schizophrenia, ranging between +2.56 years and +7.8 years (Nenadic et al., 2017; Kaufmann et al., 2019; Koutsouleris et al., 2014; Koutsouleris et al., 2015; Schnack et al., 2016; Shahab et al., 2019). Of the psychiatric disorders which have been investigated in brain age studies, MDD is perhaps the most comparable to SAD, as mood and anxiety disorders are often comorbid (Stein et al., 2017). Two related studies found

a large statistically significant positive BAG in MDD of +4 years (Koutsouleris et al., 2014; Koutsouleris et al., 2015), while Besteher et al. (2019) found a small, non-significant positive BAG of +0.41. Kaufmann et al. (2019) which is a multi-site study like the present dissertation, found small positive BAGs of +0.86 for MDD, however it was not statistically significant. Han et al. (2020), also a multi-site study, found a small, uncorrected BAG of +0.68 years. When adjusting for age, age^2 and sex, Han et al. (2020) found that MDD patients had a BAG larger than controls by +1.08 years.

The smaller BAGs observed by Besteher et al. (2019), Han et al. (2020) and Kaufmann et al. (2019) discussed above are closer in size to that found in the present study, keeping in mind that this BAG was not significantly larger in SAD patients compared to HCs in the present dissertation according to the best-fitting brain age model. However, the best-fitting model in terms of MAE may not be the most reliable model for detecting a BAG between patients and controls. Bashyam et al. (2020) assessed three levels of model fit (tight fit: MAE = 3.70, moderate fit: MAE = 5.92 and loose fit: MAE = 7.65) for a deep brain network used to estimate the brain age of individuals diagnosed with schizophrenia, MDD, mild cognitive impairment and Alzheimer's disease compared to HCs. The model with the moderate MAE had the most significant BAGs across the clinical groups and the largest effect sizes between all four disease groups and HCs. Tighter fitting models tend to focus more exclusively on features that are only affected by age, likely ignoring the effects of disease processes. However, looser fitting models may miss more subtle patterns of disease processes, in turn missing integral features of brain ageing. Based on these findings, neither the Cole nor the Kaufmann model can be declared more reliable for the present study data.

Of note, the analyses investigating whether a positive BAG is present in the SAD group and if the BAG will be significantly larger in SAD patients compared to HCs, were rerun in the Cole database (please see section 3.4 of the supplementary materials for more

details). The analyses were rerun without the participant scans from University of Chicago and University of Michigan to allow for better comparability of the results between the models, as these were the study sites that had been excluded during QC in the Kaufmann database. Significant findings were not observed; however, this is to be expected due to the smaller sample size. Therefore, the University of Chicago and the University of Michigan did not account for the lack of significant findings in this analysis.

4.2. The BAG and Clinical Characteristics

Contrary to expectations, there was no significant difference in BAG between the SAD group with psychiatric comorbidities (*n*=32, MDD: *n*=10, other anxiety disorder: *n*=24), the SAD group without psychiatric comorbidities and HCs in the Cole model. It was hypothesised that SAD patients with psychiatric comorbidities would have the largest BAG due to the evidence in the literature of a positive BAG in MDD. As discussed in chapter 4.1, a BAG of at least half a year to one year was observed in MDD patients in most of the articles included in the systematic review. Based on cortical thickness and surface area data and subcortical and cerebellar volumes, the Kaufmann model identified the comorbidity-free SAD group to have the largest BAG compared to HCs and the SAD group with secondary psychiatric comorbidities (*n*=19, MDD: *n*=9, other anxiety disorder: *n*=12). However, the comorbidity-free SAD group did not have a significantly larger BAG than the SAD group with comorbidities ($p = .749$). Of note, the sample size for the SAD group with comorbidities was small in both models (Cole model *n*=32, Kaufmann model *n*=19). This may have limited power to detect group differences in BAG between patients with and without comorbidities. Finally, the sample is not representative due to the small number of research sites included in this analysis and the variation in comorbidity exclusion criteria between the research site samples. Therefore, it is uncertain whether there is a true lack of additive effects for multiple psychiatric diagnoses.

The final analysis in the present study found no significant association between the BAG and symptom severity or medication use in both models. A positive association between BAG and symptom severity was expected as higher symptom severity may indicate a more severe disorder, in turn reflecting a more advanced brain ageing process. However, findings in the literature are mixed. One study observed a positive association between GMV in the left putamen and symptom severity (Bas-Hoogendam et al., 2017), suggesting that patients with higher symptom severity had larger GMV. Another study observed a negative association between cortical thickness in the right rostral ACC and symptom severity (Frick et al., 2013), suggesting that patients with higher symptom severity had thinner cortical thickness. Nevertheless, both models in the present dissertation were consistent in not observing an association between symptom severity and BAG. The lack of findings could be a result of the severity measure. It would perhaps be more appropriate and informative if lifetime history of symptoms had been measured rather than current symptom severity. It is also important to note that there was missing LSAS data from multiple research sites, resulting in a smaller sample size for this analysis. In addition, the biological heterogeneity of anxiety disorders could contribute to the lack of findings in this analysis as well as others. Neural circuit dysfunction and psychiatric symptomatology differ between individuals with the same psychiatric diagnosis (Boschloo et al., 2015; Williams, 2016). It is therefore possible that an increased BAG would be present in only a sub-group of patients. However, it was not possible for the present dissertation to capture all the relevant variables that could be used for subtyping SAD patients.

Most studies in the systematic review that assessed medication use were studies with psychotic or bipolar disorder patients who were taking antipsychotic medication and lithium. Han et al. (2020) found that antidepressant medication-free MDD patients and MDD patients taking antidepressant medication at the time of the scan had a significantly larger BAG than

controls. However, when compared with one another there was no significant difference in BAG between these two groups. Moreover, Kolenic et al. (2018) found that brain age scores were not associated with antidepressant, anticonvulsant or antipsychotic use, and brain age was not associated with number of medication types used. Based on this, it was hypothesised that there would be no association between the BAG and medication use in the present dissertation. The small effect sizes observed for medication use in the sub-analysis assessing this association are therefore in line with this expectation. Once again, medication use data was missing from certain research sites, resulting in a smaller sample size.

4.3. Similarities and Differences Between Model Methods

Both models used T1-weighted structural MRI data, however the Cole model estimated brain age based on VBM-volumetric data while the Kaufmann model estimated brain age based on cortical thickness and surface area data as well as subcortical and cerebellar volumes. VBM is a technique that measures the volume of the whole brain or its sub-parts and performs statistical tests across all the voxels in an image to detect group differences in volume (Whitwell, 2009). VBM uses a process called spatial normalisation which allows statistical analyses to be performed across MRI scans from multiple individuals by matching them spatially. This causes a location in one individual's MRI to correspond to the same location in another individual's MRI. An additional step in VBM is the smoothing of the data which reduces the variance between individuals and increases sensitivity to detect changes (Whitwell, 2009). On the other hand, the Kaufmann model performed surface-based morphometry and subcortical segmentation on FreeSurfer-derived cortical thickness, surface area and subcortical volume, which required less intermediate image processing steps compared to VBM. Therefore, the models estimate brain age on substantially different feature sets.

Like the Kaufmann model, Han et al. (2020) estimated brain age based on cortical thickness, surface area and cortical thickness/cerebellar volume, and observed significantly larger BAGs in their MDD groups compared to HCs. Han et al. (2020) found the strongest negative correlations between cortical thickness and predicted brain age ($r = -0.44$), followed by subcortical volume ($r = -0.34$) and surface area features ($r = -0.17$). The model depended mostly on cortical thickness features to make accurate age predictions. This finding is in line with evidence that highlights the importance of the relationship between cortical thickness and biological ageing as compared to surface area. Cortical thickness may provide more information on morphometric changes associated with ageing than surface area. Surface area may be less sensitive to these morphometric changes than other surface-based features. Additionally, it suggests that models using cortical thickness for brain age prediction are more sensitive and powerful than those using VBM, possibly because surface-based features are more sensitive to human ageing than voxel-based features (Wang et al., 2014). Like the current study, Besteher et al. (2019), Koutsouleris et al. (2014) and Koutsouleris et al. (2015) used VBM volumetric data to predict brain age. Although Koutsouleris et al. (2015) found the largest BAG in MDD, Besteher et al. (2019) observed the smallest BAG in their MDD group which was not significantly different than the HC group. Nonetheless, the Kaufmann brain age model may be able to capture ageing-related alterations in cortical thickness in SAD patients that cannot be detected through the VBM-based Cole model. However, the diagnosis-by-age interaction observed in the Cole model and the post hoc analyses show that a significantly larger BAG is detected in SAD patients, compared to HCs, when younger participants are removed from analyses, bringing the results of the two models more in line with one another.

Kaufmann et al. (2019) assessed the correlation between their main xgboost model and two additional models in their supplementary materials, one of which was Cole's

BrainageR. Kaufmann predicted brain age and Cole predicted brain age had a moderate positive relationship (r = 0.66) with one another. This was the strongest correlation between brain age models reported in Kaufmann et al. (2019). On the other hand, Kaufmann BAG and Cole BAG had a weak relationship (r = 0.32) in that particular study. The current dissertation observed a similar moderate association between Kaufmann predicted brain age and Cole predicted brain age (Pearson correlation = 0.511 ($p < .001$). Although the models differed in whether they observed a significant difference in BAG between SAD patients and HCs, they produced similar brain age estimates. However, there was no association for BAG between the two models in this dissertation (Pearson correlation = 0.120, $p = .066$). This suggests that brain age estimates produced by these two models represent substantially different outcomes.

4.4. Limitations and Strengths of the Present Study

The most notable study limitation was that the model which used cortical thickness/surface area and subcortical/cerebellar information to estimate brain age (the Kaufmann model) did not fit the study data well, as demonstrated particularly by the high MAE in the HC group (10.91 (6.7)). However, it is important to keep in mind that the tightest fitting model is not necessarily the best at identifying a BAG between patients and controls (Bashyam et al., 2020). In addition, although the sample sizes were larger than certain existing brain age studies (e.g. Besteher et al., 2019; de Nooij et al., 2019), they were perhaps still small compared to others (e.g. Han et al., 2020; Kaufmann et al., 2019). Finally, the LSAS may not have been the most appropriate measure to assess the association between the BAG and symptom severity as it only assesses an individual's SAD symptoms around the time they take the test. Nevertheless, the LSAS is recognised as the gold-standard for measuring current SAD symptom severity.

In addition to limitations, the present study also has various strengths. Brain age estimation is a novel and promising method used to explore the ageing brain in psychiatric

disease. Moreover, it may be able to provide information on the relationship between advanced brain ageing and clinical characteristics, such as disorder symptom severity. One of the most notable strengths of the current study is that a second brain age model was applied when the original model demonstrated sub-optimal fit to the data. This ensured the use of a method with better model fit to the study data, while at the same time allowing brain age to be measured using both VBM-volumetric based data and cortical thickness/ surface area data and subcortical/cerebellar volumetrics. In addition, both models used in this study employed a large sample for model training (Cole training sample *n*=3377, Kaufmann training sample *n*=35,474) that was independent from their testing samples and from the present sample.

A systematic review of the literature was conducted to improve the quality of the thesis. The systematic review provided an overview of previously conducted research which has explored the BAG in adults diagnosed with axis I psychiatric disorders compared to HCs, based on structural MRI scans. I was therefore able to identify a gap in the literature regarding the BAG in adults diagnosed with anxiety disorders. To the best of my knowledge, this is the first study to date that investigates the BAG in SAD in adults. It therefore addresses an important gap in the literature, providing insight into the ageing brain in adults diagnosed with SAD, as well as its relationship to clinical characteristics including comorbidities, symptom severity and medication use.

4.5. Recommendations for Future Research

Future research should employ multiple different brain age models that utilise different feature sets, to identify robust estimates of brain age. If there is an advanced ageing process present in SAD, a good-fitting model trained on cortical thickness/surface area and subcortical/cerebellar volumes may be better able to identify it. Where feasible, future studies should train brain age models on study-specific control groups to achieve more reliable BAG estimates.

In addition, a different measure of symptomatology should be used in future studies than the LSAS. Measuring patient lifetime history of SAD would allow for a more comprehensive measurement of disorder severity over the lifetime, rather than only assessing symptoms in the time around test administration. This could possibly identify if patients with a more severe and persistent disorder have a greater BAG compared to those with a milder disorder. Finally, future studies should try to include larger sample sizes, especially in sub-analyses where sample sizes decrease, to improve the study's power to identify group differences.

4.6. Conclusion

In the first study to explore the BAG in adults diagnosed with SAD, contradictory evidence for a larger BAG between patients with SAD than HCs was found. The contradictory results observed by the Cole model and the Kaufmann model may be due to the different information used to estimate brain age (voxel-based volumetric data and cortical thickness/surface area and subcortical/cerebellar volumetrics, respectively). No evidence was found in support of an advanced brain ageing process in patients with SAD compared to HCs based on voxel-based volumetric data. In addition, no evidence was found in support of an association between BAG and psychiatric comorbidities, symptom severity or medication use. However, the possibility of a BAG being present in SAD cannot be completely excluded. The present dissertation found evidence of advanced brain ageing in patients with SAD compared to HCs based on surface area segmentations, as utilised by the Kaufmann model. Furthermore, both models had largely overlapping confidence intervals for mean difference in BAG between SAD patients and controls. Therefore, if a positive BAG is present in adults diagnosed with SAD compared to HCs, it is likely to be small. This should be verified in future research using multiple different machine learning algorithms based on different feature sets, to acquire more robust and reliable estimates of brain age.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC.

Ashburner, J. (2010). VBM tutorial. *Wellcome Trust Centre for Neuroimaging*. London, UK.

Bas-Hoogendam, J. M. (2019). Commentary: Gray matter structural alterations in social anxiety disorder: A voxel-based meta-analysis. *Frontiers in Psychiatry*, *10*(1). https://doi.org/10.3389/fpsyt.2019.00001.

Bas-Hoogendam, J. M., Steenbergen, H. Van, Pannekoek, J. N., Fouche, J., Lochner, C., Hattingh, C. J., Cremers, H. R., Furmark, T., Månsson, K. N., Frick, A., Engman, J., Boraxbekk, C., Carlbring, P., Andersson, G., Fredrikson, M., Straube, T., Peterburs, J., Klumpp, H., Phan, K. L. … Van Der Wee, N. J. A. (2017). Voxel-based morphometry multi-center mega-analysis of brain structure in social anxiety disorder. *NeuroImage: Clinical*, *16*, 678–688. https://doi.org/10.1016/j.nicl.2017.08.001

Bashyam, V. M., Erus, G., Doshi, J., Habes, M., Nasrallah, I. M., Truelove-Hill, M., Srinivasan, D., Mamourian, L., Pomponio, R., Fan, Y., Launer, L. J., Masters, C. L., Maruff, P., Zhuo, C., Völzke, H.,Johnson, S. C., Fripp, J., Koutsouleris, N., Satterthwaite, T. D. … Davatzikos, C. (2020). MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. *Brain*, *143*(7), 2312-2324. https://doi.org/10.1093/brain/awaa160

Baur, V., Hänggi, J., Rufer, M., Delsignore, A., Jäncke, L., Herwig, U., & Brühl, A. B. (2011). White matter alterations in social anxiety disorder. *Journal of Psychiatric Research*, *45*(10), 1366-1372. https://doi.org/10.1016/j.jpsychires.2011.05.007

Beck, A.T., Steer, R.A., Carbin, M.G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*(1), 77–100. https://doi.org/10.1016/0272-7358(88)90050-5

Beesdo, K., Bittner, A., Pine, D. S., Stein, M. B., Höfler, M., Lieb, R., & Wittchen, H. U. (2007). Incidence of social anxiety disorder and the consistent risk for secondary depression in the first three decades of life. *Archives of General Psychiatry*, *64*(8), 903-912. https://doi.org/10.1001/archpsyc.64.8.903

Besteher, B., Gaser, C., & Nenadić, I. (2019). Machine-learning based brain age estimation in major depression showing no evidence of accelerated aging. *Psychiatry Research: Neuroimaging*, *290*, 1-4. https://doi.org/10.1016/j.pscychresns.2019.06.001.

Bijl, R. V., Ravelli, A., & Van Zessen, G. (1998). Prevalence of psychiatric disorder in the general population: Results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology*, *33*(12), 587-595. https://doi.org/10.1007/s001270050098

Boehme, S., Mohr, A., Becker, M.P., Miltner, W.H., & Straube, T. (2014). Area-dependent time courses of brain activation during video-induced symptom provocation in social anxiety disorder. *Biololgy of Mood and Anxiety Disorders*, *4*(6). https://doi.org/10.1186/2045-5380-4-6

Boehme, S., Ritter, V., Tefikow, S., Stangier, U., Strauss, B., Miltner, W.H.R., & Straube, T. (2015). Neural correlates of emotional interference in social anxiety disorder. *PLoS One*, *10*(6). https:/doi.org/10.1371/journal.pone.0128608

Boehme, S., Ritter, V., Tefikow, S., Stangier, U., Strauss, B., Miltner, W.H.R., & Straube, T. (2014). Brain activation during anticipatory anxiety in social anxiety disorder. *Social Cognitive and Affective Neuroscience*, *9*(9), 1413–1418. https://doi.org/10.1093/scan/nst129

Boeke, E. A., Holmes, A. J., & Phelps, E. A. (2020). Toward robust anxiety biomarkers: A machine learning approach in a large-scale sample. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(8), 799-807.

https://doi.org/10.1016/j.bpsc.2019.05.018

Boschloo, L., van Borkulo, C. D., Rhemtulla, M., Keyes, K. M., Borsboom, D., & Schoevers, R. A. (2015). The network structure of symptoms of the diagnostic and statistical manual of mental disorders. *PloS One*, *10*(9). https://doi.org/10.1371/journal.pone.0137621

Brühl, A. B., Hänggi, J., Baur, V., Rufer, M., Delsignore, A., Weidt, S., Jancke, L., & Herwig, U. (2014). Increased cortical thickness in a frontoparietal network in social anxiety disorder. *Human Brain Mapping*, *35*(7), 2966–2977. https://doi.org/10.1002/hbm.22378

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). https://doi.org/10.1145/2939672.2939785

Chung, Y., Addington, J., Bearden, C. E., Cadenhead, K., Cornblatt, B., Mathalon, D. H., McGlashan, T., Perkins, D.,Seidman, L., Tsuang, M., Walker, E., Woods, S. W., McEwan, S., van Erp, T. G. M., & Cannon, T. (2018). Use of machine learning to determine deviance in neuroanatomical maturity associated with future psychosis in youths at clinically high risk. *JAMA Psychiatry*, *75*(9), 960-968. https://doi.org/10.1001/jamapsychiatry.2018.1543

Cole, J. H. (2017). Neuroimaging-derived brain-age: An ageing biomarker? *Aging*, *9*(8), 1861–1862. https://doi.org/10.18632/aging.101286

Cole, J. H., & Franke, K. (2017). Predicting age using neuroimaging: Innovative brain ageing biomarkers. *Trends in Neurosciences*, *40*(12), 681-690. https://doi.org/10.1016/j.tins.2017.10.001

Cole, J. H., Franke, K., & Cherbuin, N. (2019). Quantification of the biological age of the brain using neuroimaging. In *Biomarkers of human aging* (pp. 293-328). Springer, Cham

Cole, J. H., Leech, R., Sharp, D. J., & Alzheimer's Disease Neuroimaging Initiative. (2015). Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Annals of Neurology*, *77*(4), 571-581. https://doi.org/10.1002/ana.24367

Cole, J. H., Marioni, R. E., Harris, S. E., & Deary, I. J. (2018). Brain age and other bodily ‘ ages ’ : Implications for neuropsychiatry. *Molecular Psychiatry*, *24*(2), 266-281. https://doi.org/10.1038/s41380-018-0098-1

Cole, J. H., Poudel, R. P. K., Tsagkrasoulis, D., & Caan, M. W. A. (2017). Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *Neuroimage*, *163*, 115-124. https://doi.org/10.1016/j.neuroimage.2017.07.059

Cole, J. H., Ritchie, S. J., Bastin, M. E., Hernández, M. V., Maniega, S. M., Royle, N., Corley, J., Pattie, A., Harris, S. E., Zhang, Q., Wray, N. R., Redmond, P., Marioni, R. E., Starr, J. M., Cox, S. R., Wardlaw, J. M., Sharp, D. J., & Deary, I. J. (2018). Brain age predicts mortality. *Molecular Psychiatry*, *23*(5), 1385-1392. https/doi.org/10.1038/mp.2017.62

Cremers, H.R., Veer, I.M., Spinhoven, P., Rombouts, S.A.R.B., & Roelofs, K. (2015). Neural sensitivity to social reward and punishment anticipation in social anxiety disorder. *Frontiers in Behavioural Neuroscience*, 8, 439. https://doi.org/10.3389/fnbeh.2014.00439

Cremers, H.R., Veer, I.M., Spinhoven, P., Rombouts, S.A.R.B., Yarkoni, T., Wager, T.D., & Roelofs, K. (2014). Altered cortical-amygdala coupling in social anxiety disorder during the anticipation of giving a public speech. *Psychological Medicine*, *45*(7), 1–9.

https://doi.org/10.1017/S0033291714002657

de Nooij, L., Harris, M. A., Hawkins, E. L., Clarke, T. K., Shen, X., Chan, S. W., Ziermans, T. B., McIntosh, A. M., & Whalley, H. C. (2019). Longitudinal trajectories of brain age in young individuals at familial risk of mood disorder. *Wellcome Open Research*, *4*(206), 206. https://doi.org/10.12688/wellcomeopenres.15617.3

Fehm, L., Pelissolo, A., Furmark, T., & Wittchen, H. (2005). Size and burden of social phobia in Europe. *European Psychopharmacology*, *15*(4), 453–462. https://doi.org/10.1016/j.euroneuro.2005.04.002

Filzmoser, P., Liebmann, B., & Varmuza, K. (2009). Repeated double cross validation. *Journal of Chemometrics: A Journal of the Chemometrics Society*, *23*(4), 160-171. https://doi.org/10.1002/cem.1225

First, M., Spitzer, R., Gibbon, M., Williams, J., & Benjamin, L. (1998). *Structured clinical interview for DSM-IV axis I disorders: SCID, version 2.0.* New York: Biometrics Research Dept, New York State Psychiatric Institute.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., & Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341-355. https://doi.org/10.1016/s0896-6273(02)00569-x

Franke, K., & Gaser, C. (2019). Ten years of BrainAGE as a neuroimaging biomarker of brain aging: What insights have we gained? *Frontiers in Neurology*, *10*, 789. https://doi.org/10.3389/fneur.2019.00789

Franke, K., Ziegler, G., Klöppel, S., Gaser, C., & Alzheimer's Disease Neuroimaging Initiative. (2010). Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various

parameters. *Neuroimage*, *50*(3), 883-892. https://doi.org/10.1016/j.neuroimage.2010.01.005

Frick, A., Gingnell, M., Marquand, A. F., Howner, K., Fischer, H., Kristiansson, M., Williams, S. C. R.,Fredriksn, M., & Furmark, T. (2014). Classifying social anxiety disorder using multivoxel pattern analyses of brain function and structure. *Behavioural Brain Research*, *259*, 330–335. https://doi.org/10.1016/j.bbr.2013.11.003

Frick, A., Howner, K., Fischer, H., Kristiansson, M., & Furmark, T. (2013). Altered fusiform connectivity during processing of fearful faces in social anxiety disorder. *Translational Psychiatry*, *3*(10), e312-6. https://doi.org/10.1038/tp.2013.85

Geiger, M.J., Domschke, K., Ipser, J., Hattingh, C., Baldwin, D.S., Lochner, C., & Stein, D.J. (2016). Altered executive control network resting-state connectivity in social anxiety disorder. T*he World Journal of Biological Psychiatry*, *17*(1), 47–57. https://doi.org/10.3109/15622975.2015.1083613

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., & Van Esse, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171-178. https://doi.org/10.1038/nature18933

Guggenmos, M., Schmack, K., Sekutowicz, M., Garbusow, M., Sebold, M., Sommer, C., Smolka., M. N., Wittchen, H., Zimmerman, U. S., Heinz, A., & Sterzer, P. (2017). Quantitative neurobiological evidence for accelerated brain aging in alcohol dependence. *Translational Psychiatry*, *7*(12), 1-7. https://doi.org/10.1038/s41398-017-0037-y

Hajek, T., Franke, K., Kolenic, M., Capkova, J., Matejka, M., Propper, L., Uher, R., Stopkova, P., Paus, T., Kopecek, M., Spaniel, F., & Alda, M. (2019). Brain age in early stages of bipolar disorders or schizophrenia. *Schizophrenia Bulletin*, *45*(1), 190-198.

https://doi.org/10.1093/schbul/sbx172

Han, L. K., Dinga, R., Hahn, T., Ching, C. R., Eyler, L. T., Aftanas, L., Aghajani, M., Aleman,A., Baune, B. T., Berger, K., Brak, I., Filho, G. B., Carballedo, A., Connolly, C. G., Couvy-Duchesne, B., Cullen, K. R., Dannlowski, U., Davey, C. G., & Dima, D. (2020). Brain aging in major depressive disorder: Results from the ENIGMA major depressive disorder working group. *Molecular Psychiatry*, 1-16. https://doi.org/10.1038/s41380-020-0754-0

Hattingh, C. J., Ipser, J., Tromp, S., Syal, S., Lochner, C., Brooks, S. J. B., & Stein, D. J. (2013). Functional magnetic resonance imaging during emotion recognition in social anxiety disorder: An activation likelihood meta-analysis. *Frontiers in Human Neuroscience*, *6*, 347. https://doi.org/10.3389/fnhum.2012.00347

Heimberg, R. G., Horner, K. J., Juster, H. R., Safren, S. A., Brown, E. J., Schneier, F. R., & Liebowitz, M. R. (1999). Psychometric properties of the Liebowitz Social Anxiety Scale. *Psychological Medicine, 29*(1), 199-212. https://doi.org/10.1017/s0033291798007879

Hoen, P. W., Rosmalen, J. G. M., Schoevers, R. A., Huzen, J., Van Der Harst, P., & De Jonge, P. (2013). Association between anxiety but not depressive disorders and leukocyte telomere length after 2 years of follow-up in a population-based sample. *Psychological Medicine*, *43*(4), 689-697. https://doi.org/10.1017/S0033291712001766

Hofmann, S. G. (2007). Cognitive factors that maintain social anxiety disorder: A comprehensive model and its treatment implications. *Cognitive Behaviour Therapy, 36*(4), 193–209. https://doi.org/10.1080/16506070701421313

Howells, F.M., Hattingh, C.J., Syal, S., Breet, E., Stein, D.J., Lochner, C., 2015. 1-H magnetic resonance spectroscopy in social anxiety disorder. *Progress in*

Neuropsychopharmacology and Biological Psychiatry, *58*, 97–104. https://doi.org/10.1016/j.pnpbp.2014.12.008

Irle, E., Barke, A., Lange, C., & Ruhleder, M. (2014). Parietal abnormalities are related to avoidance in social anxiety disorder: A study using voxel-based morphometry and manual volumetry. *Psychiatry Research: Neuroimaging*, *224*(3), 175-183. https:/doi.org/-10.1016/j.pscychresns.2014.08.013

Irle, E., Ruhleder, M., Lange, C., Seidler-brandler, U., Salzer, S., Dechent, P., Weniger, G., Leibing, E., & Leichsenring, F. (2010). Reduced amygdalar and hippocampal size in adults with generalized social phobia. *Journal of Psychiatry and Neuroscience*, *35*(2), 126–131. https://doi.org/10.1503/jpn.090041

Johnson, S. B., Blum, R. W., & Giedd, J. N. (2009). Adolescent maturity and the brain: the promise and pitfalls of neuroscience research in adolescent health policy. *Journal of Adolescent Health*, *45*(3), 216-221. https://doi.org/10.1016/j.jadohealth.2009.05.016.

Kambeitz-Ilankovic, L., Haas, S. S., Meisenzahl, E., Dwyer, D. B., Weiske, J., Peters, H., Möller, H., Falkai, P., & Koutsouleris, N. (2019). Neurocognitive and neuroanatomical maturation in the clinical high-risk states for psychosis: A pattern recognition study. *NeuroImage: Clinical*, *21*, 101624. https://doi.org/10.1016/j.nicl.2018.101624

Karatzoglou A, Smola A, Hornik K, Zeileis A (2004). Kernlab – An S4 package for kernel methods in R. *Journal of Statistical Software*, *11*(9), 1-20. http://www.jstatsoft.org/v11/i09/

Kaufmann, T., van der Meer, D., et al (2019). Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nature Neuroscience*, *22*(10), 1617-1623. https://doi.org/10.1038/s41593-019-0471-7

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey

Replication. *Archives of General Psychiatry*, *62*(6), 617-627. https://doi.org/10.1001/archpsyc.62.6.617.

Kessler, R. C., Petukhova, M., Sampson, N. A., & Zaslavsky, A. M. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, *21*(3), 169–184. https://doi.org/10.1002/mpr

Kessler, R. C., Stein, M. B., & Berglund, P. (1998). Social phobia subtypes in the National Comorbidity Survey. *American Journal of Psychiatry*, *155*(5), 613-619. https://doi.org/10.1176/ajp.155.5.613.

Kessler, R. C., & Üstün, T. B. (2004). The world mental health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). *International Journal of Methods in Psychiatric Research*, *13*(2), 93-121. https://doi.org/10.1002/mpr.168

Klumpp, H., Fitzgerald, D.A., Piejko, K., Roberts, J., Kennedy, A.E., & Phan, K.L. (2015). Prefrontal control and predictors of cognitive behavioral therapy response in social anxiety disorder. *Social Cognitive and Affective Neuroscience*, *11*(4), 630–40. https://doi.org/10.1093/scan/nsv146

Kolenic, M., Franke, K., Hlinka, J., Matejka, M., Capkova, J., Pausova, Z., Uher, H., Alda, M., Spaniel, F., & Hajek, T. (2018). Obesity, dyslipidemia and brain age in first-episode psychosis. *Journal of Psychiatric Research*, *99*, 151-158. https://doi.org/10.1016/j.jpsychires.2018.02.012

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Koutsouleris, N., Davatzikos, C., Borgwardt, S., Gaser, C., Bottlender, R., Frodl, T., Falkai,

P., Riecher-Rössler., Möller, H., Reiser, M., Pantelis, C., & Meisenzahl, E. (2014). Accelerated brain aging in schizophrenia and beyond: A neuroanatomical marker of psychiatric disorders. *Schizophrenia Bulletin*, *40*(5), 1140-1153. https://doi.org/10.1093/schbul/sbt142

Koutsouleris, N., Meisenzahl, E. M., Borgwardt, S., Riecher-Rössler, A., Frodl, T., Kambeitz, J., Köhler, Y., Falkai, P., Möller, H.,Reiser, M., & Davatzikos, C. (2015). Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. *Brain*, *138*(7), 2059-2073. https://doi.org/10.1093/brain/awv111

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, *12*(5), 535. https://doi.org/10.1038/nn.2303

Landman-Peeters, K. M., Hartman, C. A., van der Pompe, G., den Boer, J. A., Minderaa, R. B., & Ormel, J. (2005). Gender differences in the relation between social support, problems in parent-offspring communication, and depression and anxiety. *Social Science & Medicine*, *60*(11), 2549-2559. https://doi.org/10.1016/j.socscimed.2004.10.024

Le, T. T., Kuplicki, R., Yeh, H. W., Aupperle, R. L., Khalsa, S. S., Simmons, W. K., & Paulus, M. P. (2018). Effect of ibuprofen on BrainAGE: A randomized, placebo-controlled, dose-response exploratory study. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(10), 836-843. https://doi.org/10.1016/j.bpsc.2018.05.002

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., Janavs, J., & Dunbar, G. C. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. *European Psychiatry*, *12*(5), 224-231. https://doi.org/10.1016/S0924-

9338(97)83296-8

Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I disorders (SCID I) and Axis II disorders (SCID II). *Clinical Psychology & Psychotherapy*, *18*(1), 75-79. https://doi.org/10.1002/cpp.693

Machado-de-Sousa, J. P., de Lima Osório, F., Jackowski, A. P., Bressan, R. A., Chagas, M. H., Torro-Alves, N., DePaula, A., Crippa, J. A. S. & Hallak, J. E. (2014). Increased amygdalar and hippocampal volumes in young adults with social anxiety. *PloS one*, *9*(2). https://doi.org/10.1371/journal.pone.0088523

Malouff, J. M., & Schutte, N. S. (2017). A meta-analysis of the relationship between anxiety and telomere length. *Anxiety, Stress, & Coping*, *30*(3), 264-272. https://doi.org/10.1080/10615806.2016.1261286

Månsson, K. N., Carlbring, P., Frick, A., Engman, J., Olsson, C. J., Bodlund, O., Furmark, T., & Andersson, G. (2013). Altered neural correlates of affective processing after internet-delivered cognitive behavior therapy for social anxiety disorder. *Psychiatry Research: Neuroimaging*, *214*(3), 229-237. https://doi.org/10.1016/j.pscychresns.2013.08.012

Meier, S. M., Mattheisen, M., Mors, O., Mortensen, P. B., Laursen, T. M., & Penninx, B. W. (2016). Increased mortality among people with anxiety disorders: Total population study. *The British Journal of Psychiatry*, *209*(3), 216-221. https://doi.org/10.1192/bjp.bp.115.171975

Meng, Y., Lui, S., Qiu, C., Qiu, L., Lama, S., Huang, X., Feng, Y., Chunyan, Z., Gong, Q., & Zhang, W. (2013). Neuroanatomical deficits in drug-naïve adult patients with generalized social anxiety disorder: A voxel-based morphometry study. *Psychiatry Research - Neuroimaging*, *214*(1), 9–15. https://doi.org/10.1016/j.pscychresns.2013.06.002

Merikangas K., Avenevoli S., Costello J., Koretz D., & Kessler R.C. (2009). National comorbidity survey replication adolescent supplement (NCS-A): I. Background and measures. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(4), 367–369. https://doi.org/10.1097/CHI.0b013e31819996f1

Moher D, Liberati A, Tetzlaff J, Altman DG, & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed1000097

Nenadić, I., Dietzek, M., Langbein, K., Sauer, H., & Gaser, C. (2017). BrainAGE score indicates accelerated brain aging in schizophrenia, but not bipolar disorder. *Psychiatry Research: Neuroimaging, 266*, 86-89. https://doi.org/10.1016/j.pscychresns.2017.05.006

Nenadic, I., Maitra, R., Langbein, K., Dietzek, M., Lorenz, C., Smesny, S., Reichenbach, J. R., Sauer, H., & Gaser, C. (2015). Brain structure in schizophrenia vs. psychotic bipolar I disorder: A VBM study. *Schizophrenia Research, 165*(2-3), 212-219. https://doi.org/10.1016/j.schres.2015.04.007

Nitschke, J. B., & Heller, W. (2005). Distinguishing neural substrates of heterogeneity among anxiety disorders. *International Review of Neurobiology, 67*, 1-42. https://doi.org/10.1016/S0074-7742(05)67001-8

Niu, X., Zhang, F., Kounios, J., & Liang, H. (2020). Improved prediction of brain age using multimodal neuroimaging data. *Human Brain Mapping, 41*(6), 1626-1643. https://doi.org/10.1002/hbm.24899

Pannekoek, J.N., van der Werff, S.J.A., van Tol, M.J., Veltman, D.J., Aleman, A., Zitman, F.G., Rombouts, S.A.R.B., & van der Wee, N.J.A. (2015). Investigating distinct and common abnormalities of resting-state functional connectivity in depression, anxiety, and their comorbid states. *European Neuropsychopharmacology, 25*(11), 1933-1942. https://doi.org/10.1016/j.euroneuro.2015.08.002

Penninx, B.W.J.H., Beekman, A.T.F., Smit, J.H., Zitman, F.G., Nolen, W.A., Spinhoven, P., Cuijpers, P., De Jong, P.J., Van Marwijk, H.W.J., Assendelft, W.J.J., Van Der Meer, K., Verhaak, P., Wensing, M., De Graaf, R., Hoogendijk, W.J., Ormel, J., & Van Dyck, R. (2008). The Netherlands Study of Depression and Anxiety (NESDA): Rationale, objectives and methods. *International Journal of Methods in Psychiatric Research*, *17*(3), 121–140. https://doi.org/10.1002/mpr.256

Perna, G., Iannone, G., Alciati, A., & Caldirola, D. (2016). Are anxiety disorders associated with accelerated aging? A focus on neuroprogression. *Neural Plasticity*. https://doi.org/10.1155/2016/8457612

Phan, K.L., Coccaro, E.F., Angstadt, M., Kreger, K.J., Mayberg, H.S., Liberzon, I., & Stein, M.B. (2013). Corticolimbic brain reactivity to social signals of threat before and after sertraline treatment in generalized social phobia. *Biological Psychiatry*, *73*(4), 329–36. https://doi.org/10.1016/j.biopsych.2012.10.003

Rehm, J., & Shield, K. D. (2019). Global burden of disease and the impact of mental and addictive disorders. *Current Psychiatry Reports*, *21*(2), 10. https//doi.org/10.1007/s11920-019-0997-0.

Riecher-Rössler, A., Aston, J., Ventura, J., Merlo, M., Borgwardt, S., Gschwandtner, U., & Stieglitz, R. D. (2008). The Basel Screening Instrument for Psychosis (BSIP): Development, structure, reliability and validity. *Fortschritte der Neurologie-Psychiatrie*, *76*(4), 207-16. https://doi.org/10.1055/s-2008-1038155

Robins, L. N., Wing, J., Wittchen, H. U., Helzer, J. E., Babor, T. F., Burke, J., Farmer, A., Jablenski, A., Pickens, R., Regier, D., Sartorius, N., & Towle, L. H. (1988). The Composite International Diagnostic Interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry*, *45*(12), 1069-1077.

https:/doi.org/10.1001/archpsyc.1988.01800360017003

Rosen, A. F., Roalf, D. R., Ruparel, K., Blake, J., Seelaus, K., Villa, L. P., Ciric, R., Cook, P. A., Davatzikos, C., Elliott, M., de La Garza, A. G., Gennatas, E. D., Quarmley, M., Schmitt, J. E., Shinohara, R. T., Tisdall, M. D., Craddock, M. C., Raquel, E. G., Gur, R. C., & Satterhwaite, T. D. (2018). Quantitative assessment of structural image quality. *Neuroimage, 169*, 407-418. https://doi.org/10.1016/j.neuroimage.2017.12.059

Ruscio, A. M., Brown, T. A., Chiu, W. T., Sareen, J., Stein, M. B., & Kessler, R. C. (2008). Social fears and social phobia in the United States: Results from the National Comorbidity Survey Replication. *Psychological Medicine, 38*(1), 15-28. https://doi.org/10.1017/S0033291707001699

Schnack, H. G., Van Haren, N. E., Nieuwenhuis, M., Hulshoff Pol, H. E., Cahn, W., & Kahn, R. S. (2016). Accelerated brain aging in schizophrenia: A longitudinal pattern recognition study. *American Journal of Psychiatry*, *173*(6), 607-616. https://doi.org/10.1176/appi.ajp.2015.15070922

Shahab, S., Mulsant, B. H., Levesque, M. L., Calarco, N., Nazeri, A., Wheeler, A. L., Foussias, G., Rajji, T. K., & Voineskos, A. N. (2019). Brain structure, cognition, and brain age in schizophrenia, bipolar disorder, and healthy controls. *Neuropsychopharmacology*, *44*(5), 898-906. https://doi.org/10.1038/s41386-018-0298-z

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, *6*(3), 309-315. https://doi.org/10.1038/nn1008

Stein, D. J., Lim, C. C. W., Roest, A. M., Jonge, P. De, Aguilar-gaxiola, S., Al-hamzawi, A., Alonso, J., Benjet, C., Bromet, E. J., Bruffaerts, R., Girolamo, G. De., Florescu, S., Gureje, O., Haro, J. M., Harris, M. G., He, Y., Hinkov, H., Horiguchi, I., Hu, C. …

WHO World Mental Health Survey Collaborators. (2017). The cross-national epidemiology of social anxiety disorder: Data from the World Mental Health Survey Initiative. *BMC Medicine*, *15*(1), 1–21. https://doi.org/10.1186/s12916-017-0889-2

Stein, J. L., Medland, S. E., Vasquez, A. A., Hibar, D. P., Senstad, R. E., Winkler, A. M., Toro, R., Appel, K., Bartecek, R., Bergman, O., Bernard, M., Brown, A. A., Cannon, D. M., Chakravarty, M. M., Christoforou, A., Domin, M., Grimm, O., Hollinshead, M., Holmes, A. J. …Thompson, P. M. (2012). Identification of common variants associated with human hippocampal and intracranial volumes. *Nature Genetics*, *44*(5), 552-561. https://doi.org/10.1038/ng.2250

Stein, M. B., & Stein, D. J. (2008). Social anxiety disorder. *Lancet*, *371*(9618) 1115–1125. https://doi.org/10.1016/S0140-6736(08)60488-2

Syal, S., Hattingh, C. J., Fouché, J. P., Spottiswoode, B., Carey, P. D., Lochner, C., & Stein, D. J. (2012). Grey matter abnormalities in social anxiety disorder: A pilot study. *Metabolic Brain Disease*, *27*(3), 299-309. https://doi.org/10.1007/s11011-012-9299-5

Taki, Y., Kinomura, S., Sato, K., Goto, R., Kawashima, R., & Fukuda, H. (2011). A longitudinal study of gray matter volume decline with age and modifying factors. *Neurobiology of Aging*, *32*(5), 907-915. https://doi.org/10.1016/j.neurobiolaging.2009.05.003

Talati, A., Pantazatos, S. P., Schneier, F. R., Weissman, M. M., & Hirsch, J. (2013). Gray matter abnormalities in social anxiety disorder: Primary, replication, and specificity studies. *Biological Psychiatry*, *73*(1), 75–84. https://doi.org/10.1016/j.biopsych.2012.05.022

Terry, R. D., DeTeresa, R., & Hansen, L. A. (1987). Neocortical cell counts in normal human adult aging. *Annals of Neurology: Official Journal of the American Neurological*

Association and the Child Neurology Society, *21*(6), 530-539. https://doi.org/10.1002/ana.410210603

Tsaluchidu, S., Cocchi, M., Tonello, L., & Puri, B. K. (2008). Fatty acids and oxidative stress in psychiatric disorders. *BMC Psychiatry*, *8*(1), 1-3. https://doi:10.1186/1471-244X-8-S1-S5

Turk, C. L., Heimberg, R. G., Orsillo, S. M., Holt, C. S., Gitow, A., Street, L. L., Schneier, F. R., & Liebowitz, M. R. (1998). An investigation of gender differences in social phobia. *Journal of Anxiety Disorders*, *12*(3), 209-223. https://doi.org/10.1016/s0887-6185(98)00010-3

Van Gestel, H., Franke, K., Petite, J., Slaney, C., Garnham, J., Helmick, C., Johnson, K., Uher, R., Alda, M., & Hajek, T. (2019). Brain age in bipolar disorders: Effects of lithium treatment. *Australian & New Zealand Journal of Psychiatry*, *53*(12), 1179-1188. https://doi.org/10.1177/0004867419857814

van Tol, M.-J., van der Wee, N.J.A., van den Heuvel, O.A., Nielen, M.M.A., Demenescu, L.R., Aleman, A., Renken, R., van Buchem, M.A., Zitman, F.G., & Veltman, D.J. (2010). Regional brain volume in depression and anxiety disorders. *Archives of General Psychiatry*, *67*(10), 1002–11. https://doi.org/10.1001/archgenpsychiatry.2010.121

Verhoeven, J. E., Revesz, D., van Oppen, P., Epel, E. S., Wolkowitz, O. M., & Penninx, B. W. (2015). Anxiety disorders and accelerated cellular ageing. *The British Journal of Psychiatry*, *206*(5), 371-378. https://doi.org/10.1192/bjp.bp.114.151027

Wang, X., Cheng, B., Luo, Q., Qiu, L., Wang, S., & Wang, S. (2018). Gray matter structural alterations in social anxiety disorder: A voxel-based meta-analysis. *Frontiers in Psychiatry*, *9*(449). https://doi.org/10.3389/fpsyt.2018.00449

Wang, J., Li, W., Miao, W., Dai, D., Hua, J., & He, H. (2014). Age estimation using cortical surface pattern combining thickness with curvatures. *Medical & Biological*

Engineering & Computing, *52*(4), 331-341. https://doi.org/10.1007/s11517-013-1131-9

Wertz, J., Caspi, A., Ambler, A., Broadbent, J., Hancox, R. J., Harrington, H., Hogan, S., Houts, R. M., Leung, J. H., Poulton, R., Purdy, S. C., Ramrakha, S., Rasmussen, L. J. H., Richmond-Rakerd, L. S., Thorne, P. R., Wilson, G. A., & Moffitt, T. E. (2021). Association of history of psychopathology with accelerated aging at midlife. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2020.4626

Whitwell, J. L. (2009). Voxel-based morphometry: An automated technique for assessing structural changes in the brain. *Journal of Neuroscience*, *29*(31), 9661-9664. https://doi.org/10.1523/JNEUROSCI.2160-09.2009

Williams, L. M. (2016). Precision psychiatry: A neural circuit taxonomy for depression and anxiety. *The Lancet Psychiatry*, *3*(5), 472-480. https://doi.org/10.1016/S2215-0366(15)00579-9

Wittchen, H. U., Stein, M. B., & Kessler, R. C. (1999). Social fears and social phobia in a community sample of adolescents and young adults: Prevalence, risk factors and co-morbidity. *Psychological Medicine*, *29*(2), 309-323. https://doi.org/10.1017/s0033291798008174.

World Health Organisation (1993). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organisation.

World Medical Association. (2013). World Medical Association Declaration of Helsinki ethical principles for medical research involving human subjects. *JAMA: Journal of the American Medical Association*, *310*(20), 2191-2194. https://doi.org/10.1001/jama.2013.281053

Zhao, Y., Chen, L., Zhang, W., Xiao, Y., Shah, C., Zhu, H., Yuan, M., Sun, H., Yue, Q., Jia,

Z., Zhang, W., Kuang, W., Gong, Q., & Lui, S. (2017). Gray matter abnormalities in non-comorbid medication-naive patients with major depressive disorder or social anxiety disorder. *EBioMedicine*, *21*, 228–235. https://doi.org/10.1016/j.ebiom.2017.06.013

Supplementary Results

1. Inclusion and Exclusion Criteria

Inclusion and exclusion criteria for the social anxiety disorder and healthy control groups are described in the below table for each study site.

Table 1.
SAD and HC Inclusion and Exclusion Criteria

Site	Inclusion criteria	Exclusion criteria
Germany: University of Jena, University of Münster[1]		Both groups: general fMRI contraindications
	SAD: primary diagnosis of SAD	SAD: diagnosis of OCD, current substance or alcohol abuse, any psychotic disorder, dementia and current primary or secondary MDD; history of seizures, head injury with LOC; a severe uncontrollable medical condition; or psychotropic medication use within the last 6 months.
	HC: healthy adults, matched to patients on age, gender and education	HC: any psychopathology
The Netherlands (NESDA): VU Medical Centre Amsterdam, University of Groningen, Leiden University Medical Centre[2]		Both groups: presence/history of major internal or neurological disorder, recent abuse (past year) or dependence on alcohol and/or drugs, hypertension, and general MRI contraindications.
	SAD: half-year diagnosis of SAD	SAD: personality disorders; axis-I disorders other than MDD, PD, SAD, or GAD; use of psychotropic medication except infrequent benzodiazepine use or stable use of SSRIs
	HC: no presence/history of depressive or anxiety disorders or any other axis-I disorder; and no psychotropic drug use	HC: lifetime DSM-IV diagnosis
The Netherlands (Social Anxiety Study): Leiden University Medical Centre[3]	SAD: primary DSM-IV diagnosis of general SAD	SAD: co-morbid anxiety, psychotic or substance abuse disorders
	HC: age, gender and education matched to SAD; no history of psychiatric diseases or use of psychotropic medication	
South Africa: University of Cape Town, Stellenbosch University[4]	SAD: primary diagnosis of SAD; right-handed.	SAD: clinically significant comorbidity; psychotic disorder; psychotropic medication use
	HC: no psychiatric disorders; right-handed.	HC: psychotic disorder and medication use
Sweden: Umeå University[5]	Both groups: age 18 or older	Both groups: diagnosed with MDD, diagnosed with a neurological, neuropsychiatric or

		any other severe psychiatric disorder, fMRI safety issue
	SAD: primary DSM-IV diagnosis of SAD	SAD: Not diagnosed as SAD according to the SPSQ, MADRS-S total score > 30, if currently on pharmacotherapy (e.g.SSRI) the dose needed to be stable at the time of the treatment, no other present psychological treatment
USA: University of Illinois[6]	Both groups: age 18-55; subject able to give informed consent; physically healthy	Both groups: clinically significant and active neurological or medical condition; primary comorbid anxiety disorder; history of bipolar disorder, schizophrenia, or presence of an organic mental syndrome, mental retardation, or pervasive developmental disorder; history/current psychotic symptoms; current alcohol/drug abuse or dependence in the past 6 months; current/recent (6 months) suicidal ideation/attempt; chronic self-injurious behaviour in the past six months; prior cognitive behavioural therapy; ongoing psychotherapy; current treatment with any psychotropic medication; MRI contraindications
	SAD: current DSM-IV SAD diagnosis; LSAS score of ≥ 55 HC: no history of Axis I or Axis II disorders	
USA: University of Chicago, University of Michigan[7]	Both groups: age 18-55; subject able to give informed consent; physically healthy	Both groups: clinically significant neurologic or medical condition; lifetime history of schizophrenia, bipolar disorder, or presence of an organic mental syndrome, pervasive development disorder or mental retardation; positive drug screen; pregnancy or lactating; left-handed; ferrous-containing metal present in the body; unable to endure small, enclosed spaces
	SAD: current DSM-IV SAD diagnosis, generalised sub-type; LSAS score of > 60	SAD: a primary comorbid anxiety disorder; current MDD or MDD within the last 6 months; HAM-D < 18; current drug or alcohol dependence or abuse, or within the last year; current suicidal ideation; diagnosis of axis II personality disorders; concomitant treatment with psychoactive or psychotropic medications within the last 2 weeks, treatment of fluoxetine

	within the last 8 weeks, and MAOIs in the last 4 weeks
HC: no history of Axis I or Axis II disorders	

Note. DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition; fMRI, functional magnetic resonance imaging; GAD, generalised anxiety disorder; HAM-D, Hamilton Depression Rating Scale; HC, healthy controls; LOC, loss of consciousness; LSAS, Liebowitz Social Anxiety Scale; MADRS-S, Montgomery-Asberg Depression Rating Scale – Self Assessment; MRI, magnetic resonance imaging; MAOIs, monoamine oxidase inhibitors; MDD, major depressive disorder; OCD, obsessive compulsive disorder; PD, panic disorder; PET, positron emission tomography; SAD, social anxiety disorder; SPSQ, Symptoms of Post-Concussion Syndrome Questionnaire; SSRIs, Selective serotonin reuptake inhibitors.
[1] Boehme et al., 2015, 2014a, 2014b. [2] Pannekoek et al., 2015; Penninx et al., 2008, 2013; van Tol et al., 2010. [3]Cremers et al., 2015, 2014. [4]Geiger et al., 2016; Hattingh et al., 2013; Howells et al., 2015; Syal et al., 2012. [5] Månsson et al., 2013. [6]Klumpp et al., 2015. [7]Phan et al., 2013.

2. Quality Control

2.1. QC Participant Exclusion

Table 2.1.1.
FreeSurfer QC Participant Exclusion, Kaufmann Model

Reason for exclusion	**No.**
FreeSurfer failed to output labels	124
General segmentation failure	6
7+ flags	14
7+ outlier regions	3
Total participants excluded	147

Table 2.1.2.
SPM QC Participant Exclusion, Cole Model

Reason for exclusion	**No.**
Misclassification	34
Thickness	32
Folding	8
Meninges	14
General segmentation failure	1
Misregistration	0
Total participants excluded	44

Note. Number of reasons for exclusion and total participants excluded do not add up as participants were usually excluded due to multiple reasons.

2.2. QC Inter-rater Reliability, Kaufmann Model

Two-way mixed, average measure, absolute agreement, interclass correlation coefficients models were calculated using SPSS v25 to assess inter-rater reliability for each brain region. This model was chosen as it assesses the degree to which the raters assign the same score to the same subject (absolute agreement) and the results are specific to the raters involved (Koo & Li, 2016). This resulted in a very good inter-rater reliability, with an overall interclass correlation coefficient value of 0.816 (CI 0.782 to 0.844; kappa = 0.604). Next, we

further checked additional MRI slices of participants with relevant QC issues in accordance with the ENIGMA protocol, to confirm data quality.

3. Descriptive Statistics per Research Site

3.1. Kaufmann Model

Table 3.1.1.
Participant Demographic and Clinical Information: NESDA - Amsterdam

	All participants	**SAD (*n*=5)**	**HC (*n*=20)**
Female n (%)	17 (68%)	2 (40%)	15 (75%)
Mean (sd) age	38.88 (9.59)	34.00 (7.18)	40.10 (9.87)
Education n (%)			
Low (5-8y)	0 (0%)	0 (0%)	0 (0%)
Intermediate (9-14y)	9 (36%)	2 (40%)	7 (35%)
High (15y+)	16 (64%)	3 (60%)	13 (65%)
Right-handed n (%)	22 (88%)	5 (100%)	17 (85%)
*Median (IQR) AOO	6 (-)	6 (-)	N/A
Comorbidity (n)	0	0	N/A
Medication use n (%)	3 (12%)	3 (60%)	N/A
*Mean (sd) FQ – Social Phobia total score	7.13 (7.46)	18.00 (14.14)	5.57 (5.29)
Mean (sd) IDS	5.42 (5.82)	13.60 (7.44)	3.26 (2.75)
Mean (sd) MADRS	2.04 (4.18)	7 (6.89)	0.74 (1.76)

Note. AOO, age of onset; FQ, Fear Questionnaire; IDS, Inventory of Depressive Symptomology; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. *AOO was missing for 2 SAD participants. BAI was not recorded in 1 SAD participant, FQ – Social Phobia total score were not recorded in 3 SAD participants and 6 HC participants, MADRS was not recorded in 1 HC participant.

Table 3.1.2.
Participant Demographic and Clinical Information: NESDA - Leiden

	All participants	**SAD (*n*=6)**	**HC (*n*=23)**
Female n (%)	21 (72.41%)	4 (66.67%)	17 (73.91%)
Mean (sd) age	39.79 (9.24)	36.50 (10.65)	40.65 (8.89)
Education n (%)			
Low (5-8y)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Intermediate (9-14y)	15 (51.70%)	4 (66.70%)	11 (47.80%)
High (15y+)	14 (48.30%)	2 (33.30%)	12 (52.20%)
Right-handed n (%)	28 (96.55%)	6 (100%)	22 (95.65%)
Median (IQR) AOO	10.50 (12)	10.50 (12)	N/A
Comorbidity (n)	0	0	N/A
Medication use n (%)	2 (6.90%)	2 (33.33%)	N/A
*Mean (sd) FQ – Social Phobia total score	8.54 (9.49)	24.83 (4.26)	4.09 (3.85)
Mean (sd) IDS	8 (11.50)	25.17 (14.83)	3.52 (4.12)
Mean (sd) MADRS	3 (7.68)	12.83 (13.67)	0.43 (1.12)

Note. AOO, age of onset; BDI, Beck Depression Inventory; FQ, Fear Questionnaire; IDS, Inventory of Depressive Symptomology; LSAS, Liebowitz Social Anxiety Scale; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. *FQ was not recorded in one HC participant.

Table 3.1.3.
Participant Demographic and Clinical Information: NESDA - Groningen

	All participants	**SAD (*n*=7)**	**HC (*n*=10)**

Female n (%)	11 (64.71%)	5 (71.43%)	6 (60.00%)
Mean (sd) age	39.35 (10.72)	35.71 (9.41)	41.90 (11.30)
Education n (%)			
Low (5-8y)	2 (11.80%)	1 (14.30%)	1 (10.00%)
Intermediate (9-14y)	5 (29.40%)	4 (57.10%)	1 (10.00%)
High (15y+)	10 (58.80%)	2 (28.60%)	8 (80.00%)
Right-handed n (%)	17 (100.00%)	7 (100.00%)	10 (100.00%)
Median (IQR) AOO	14 (16)	14 (16)	N/A
Comorbidity (n)	1	1	N/A
MDD (n)	0	0	
Anxiety (n)	1	1	
Medication use n (%)	1 (5.88%)	1 (14.29%)	N/A
*Mean (sd) FQ – Social Phobia total score	11.80 (9.14)	21.33 (3.33)	5.44 (5.05)
Mean (sd) IDS	9.12 (10.65)	17.71 (12.04)	3.10 (2.73)
Mean (sd) MADRS	4.12 (5.40)	7.86 (7.20)	1.50 (1.90)

Note. AOO, age of onset; FQ, Fear Questionnaire; IDS, Inventory of Depressive Symptomology; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. *FQ was not recorded for 2 participants (1 SAD, 1 HC).

Table 3.1.4.
Participant Demographic and Clinical Information: Henk - Leiden

	All participants	**SAD (*n*=19)**	**HC (*n*=19)**
Female n (%)	16 (42.11%)	8 (42.11%)	8 (42.11%)
Mean (sd) age	28.39 (7.91)	29.05 (8.02)	27.74 (7.96)
Education n (%)			
Low (5-8y)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Intermediate (9-14y)	6 (15.80%)	4 (21.20%)	2 (10.50%)
High (15+)	32 (84.20%)	15 (78.90%)	17 (89.50%)
Right-handed n (%)	38 (100.00%)	19 (100.00%)	19 (100.00%)
Comorbidity (n)	2	2	N/A
MDD (n)	2	2	
Anxiety (n)	0	0	
Medication use n (%)	2 (5.26%)	2 (10.53%)	N/A
Mean (sd) LSAS	53.50 (35.15)	85.53 (13.86)	21.47 (13.48)
Mean (sd) BDI	12.26 (9.95)	19.26 (9.35)	5.26 (3.54)

Note. AOO, age of onset; BDI, Beck Depression Inventory; LSAS, Liebowitz Social Anxiety Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old

Table 3.1.5.
Participant Demographic and Clinical Information: Cubic

	All participants	**SAD (*n*=11)**	**HC (*n*=11)**
Female n (%)	10 (45.45%)	6 (54.55%)	4 (36.36%)
Mean (sd) age	28.68 (7.34)	28.45 (7.75)	28.91 (7.27)
*Education n (%)			
Low (5-8y)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Intermediate (9-14y)	7 (50.00%)	5 (62.50%)	2 (33.33%)
High (15+)	7 (50.00%)	3 (37.50%)	4 (66.67%)
Right-handed n (%)	22 (100.00%)	11 (100.00%)	11 (100.00%)
*Median (IQR) AOO	16 (10)	16 (10)	N/A
Comorbidity (n)	1	1	N/A
MDD (n)	0	0	
Anxiety (n)	1	1	
*Medication use n (%)	1 (5.56%)	1 (14.29%)	N/A
*Mean (sd) LSAS	57.79 (42.56)	86.13 (27.31)	20.00 (25.84)

Note. AOO, age of onset; LSAS, Liebowitz Social Anxiety Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. *Education and LSAS were not recorded in all SAD (n = 3) and HC (n = 5) participants. AOO and Medication use was unknown in 4 SAD participants.

Table 3.1.6.
Participant Demographic and Clinical Information: Münster

	All participants	**SAD (n=17)**	**HC (n=21)**
Female n (%)	13 (34.21%)	7 (41.18%)	6 (28.57%)
Mean (sd) age	33.21 (11.00)	34.76 (13.02)	31.95 (9.2)
Right-handed n (%)	38 (100%)	17 (100%)	21 (100%)
Comorbidity (n)	9	9	N/A
MDD (n)	6	6	
Anxiety (n)	5	5	
Medication use n (%)	0 (0%)	0 (0%)	N/A
*Mean (sd) LSAS	41.55 (27.34)	66.38 (15.34)	18.18 (11.02)
*Mean (sd) BDI	9.03 (8.53)	13.88 (9.34)	4.47 (4.26)

Note. BDI, Beck Depression Inventory; LSAS, Liebowitz Social Anxiety Scale; MADRS, MDD, major depressive disorder; n, number; sd, standard deviation. MDD and Anxiety don't add up to total comorbidity due to participants having both comorbidities. *BDI and LSAS were not recorded for 1 SAD participant and 4 HCs.

Table 3.1.7.
Participant Demographic and Clinical Information: Umea

	All participants	**SAD (*n*=24)**	**HC (*n*=23)**
Female n (%)	37 (78.72%)	21 (87.50%)	16 (69.57%)
Mean (sd) age	32.21 (10.01)	32.17 (9.54)	32.26 (10.69)
Education n (%)			
Low (5-8y)	1 (2.10%)	0 (0.00%)	1 (4.30%)
Intermediate (9-14y)	10 (21.30%)	5 (20.80%)	5 (21.70%)
High (15+)	36 (76.60%)	19 (79.20%)	17 (73.90%)
Right-handed n (%)	47 (100.00%)	24 (100.00%)	23 (100.00%)
Median (IQR) AOO	15 (9)	15 (9)	N/A
Comorbidity (n)	0	0	N/A
*Medication use n (%)	13 (28.26%)	13 (56.52%)	N/A
Mean (sd) LSAS	49.26 (32.67)	77 (18.97)	20.30 (11.82)
Mean (sd) MADRS	11.06 (7.43)	15.75 (6.76)	6.17 (4.33)

Note. AOO, age of onset; LSAS, Liebowitz Social Anxiety Scale; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. *Medication status is unknown for one SAD patient.

Table 3.1.8.
Participant Demographic and Clinical Information: University of Illinois

	All participants	**SAD (*n*=11)**	**HC (*n*=11)**
Female n (%)	15 (68.18%)	8 (72.73%)	7 (63.64%)
Mean (sd) age	30.18 (9.64)	27.64 (7.45)	32.73 (11.20)
Education n (%)			
Low (5-8y)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Intermediate (9-14y)	7 (31.80%)	2 (18.20%)	5 (45.50%)
High (15+)	15 (68.20%)	9 (81.10%)	6 (54.50%)
Right-handed n (%)	20 (90.91%)	10 (90.91%)	10 (90.91%)
Median (IQR) AOO	14 (13)	14 (13)	N/A
Comorbidity (n)	6	6	N/A
MDD (n)	1	1	
Anxiety (n)	5	5	
Medication use n (%)	0 (0.00%)	0 (0.00%)	N/A
Mean (sd) LSAS	45.14 (38.52)	80.27 (17.47)	10 (9.76)
Mean (sd) BDI	8.32 (10.33)	15.55 (10.32)	1.09 (1.64)

Note. AOO, age of onset; BDI, Beck Depression Inventory; LSAS, Liebowitz Social Anxiety Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old

3.2. Cole Model

Table 3.2.1.
Participant Demographic and Clinical Information: NESDA - Amsterdam

	All participants	**SAD (*n*=6)**	**HC (*n*=24)**
Female n (%)	17 (56.70%)	2 (33.30%)	15 (62.50%)
Mean (sd) age	38.73 (9.41)	36.17 (8.33)	39.38 (9.71)
Education n (%)			
Low (5-8y)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Intermediate (9-14y)	12 (40.00%)	3 (50.00%)	9 (37.50%)
High (15y+)	18 (60.00%)	3 (50.00%)	15 (62.50%)
Right-handed n (%)	27 (90.00%)	6 (100.00%)	21 (87.50%)
*Median (IQR) AOO	8 (9)	8 (9)	N/A
Comorbidity (n)	1	1	N/A
MDD (n)	0	0	
Anxiety (n)	1	1	
Medication use n (%)	4 (13.33%)	4 (66.67%)	N/A
*Mean (sd) FQ – Social Phobia total score	8 (7.56)	20 (10.58)	6 (4.95)
Mean (sd) IDS	6.93 (9.61)	19.33 (15.54)	3.7 (3.11)
*Mean (sd) MADRS	2.83 (5.75)	10 (9.59)	0.96 (1.85)

Note. AOO, age of onset; FQ, Fear Questionnaire; IDS, Inventory of Depressive Symptomology; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. *AOO was missing for 2 SAD participants. FQ was not recorded in 3 SAD participants and 6 HCs. MADRS was not recorded in 1 HC participant.

Table 3.2.2.
Participant Demographic and Clinical Information: NESDA - Leiden

	All participants	**SAD (*n*=7)**	**HC (*n*=24)**
Female n (%)	24 (77.40%)	5 (71.40%)	19 (79.20%)
Mean (sd) age	39.90 (10.03)	38.86 (11.55)	40.21 (9.80)
Education n (%)			
Low (5-8y)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Intermediate (9-14y)	15 (48.40%)	5 (71.40%)	10 (41.70%)
High (15y+)	16 (51.60%)	2 (28.60%)	14 (58.30%)
Right-handed n (%)	30 (96.80%)	7 (100.00%)	23 (95.80%)
Median (IQR) AOO	14 (17)	14 (17)	N/A
Comorbidity (n)	0	0	N/A
Medication use n (%)	2 (6.45%0	2 (28.57%)	N/A
*Mean (sd) FQ – Social Phobia total score	8 (7.56)	25.86 (4.74)	4.04 (4.29)
Mean (sd) IDS	6.93 (9.61)	27.29 (14.65)	4.04 (4.29)
Mean (sd) MADRS	3.65 (8.50)	14.57 (13.30)	0.46 (1.10)

Note. AOO, age of onset; BDI, Beck Depression Inventory; FQ, Fear Questionnaire; IDS, Inventory of Depressive Symptomology; LSAS, Liebowitz Social Anxiety Scale; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. * FQ was not recorded in one HC participant.

Table 3.2.3.
Participant Demographic and Clinical Information: NESDA - Groningen

	All participants	**SAD (*n*=7)**	**HC (*n*=11)**
Female n (%)	11 (61.10%)	5 (71.40%)	6 (54.50%)
Mean (sd) age	39.67 (10.48)	35.71 (9.41)	42.18 (10.76)
Education n (%)			
Low (5-8y)	2 (11.10%)	1 (14.30%)	1 (9.10%)
Intermediate (9-14y)	5 (27.80%)	4 (57.10%)	1 (9.10%)

High (15y+)	11 (61.10%)	2 (28.60%)	9 (81.80%)
Right-handed n (%)	18 (100.00%)	7 (100.00%)	11 (100.00%)
Median (IQR) AOO	14 (16)	14 (16)	N/A
Comorbidity (n)	1	1	N/A
MDD (n)	0	0	
Anxiety (n)	1	1	
Medication use n (%)	1 (5.56%)	1 (14.29%)	N/A
*Mean (sd) FQ – Social Phobia total score	11.06 (9.31)	21.33 (3.33)	4.9 (5.07)
Mean (sd) IDS	9 (10.35)	17.71 (12.04)	3.45 (2.84)
Mean (sd) MADRS	4.06 (5.48)	7.86 (7.20)	1.64 (1.86)

Note. AOO, age of onset; FQ, Fear Questionnaire; IDS, Inventory of Depressive Symptomology; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. FQ was not recorded for 2 participants (1 SAD, 1 HC).

Table 3.2.4.
Participant Demographic and Clinical Information: Henk - Leiden

	All participants	**SAD (*n*=20)**	**HC (*n*=19)**
Female n (%)	17 (43.60%)	9 (45.00%)	8 (42.10%)
Mean (sd) age	28.31 (7.83)	28.85 (7.86)	27.74 (7.96)
Education n (%)			
Low (5-8y)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Intermediate (9-14y)	6 (15.40%)	4 (20.00%)	2 (10.50%)
High (15y+)	33 (84.60%)	16 (80.00%)	17 (89.50%)
Right-handed n (%)	39 (100.00%)	20 (100.00%)	19 (100.00%)
Comorbidity (n)	2	2	N/A
MDD (n)	2	2	
Anxiety (n)	0	0	
Medication use n (%)	2 (5.13%)	2 (10.00%)	N/A
Mean (sd) LSAS	54.38 (35.12)	85.65 (13.50)	21.47 (13.48)
Mean (sd) BDI	12.36 (9.83)	19.10 (9.13)	5.26 (3.54)

Note. BDI, Beck Depression Inventory; LSAS, Liebowitz Social Anxiety Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old

Table 3.2.5.
Participant Demographic and Clinical Information: Cubic

	All participants	**SAD (*n*=11)**	**HC (*n*=11)**
Female n (%)	10 (45.50%)	6 (54.50%)	4 (36.40%)
Mean (sd) age	28.68 (7.34)	28.45 (7.75)	28.91 (7.27)
*Education n (%)			
Low (5-8y)	0 (0.00%)	0 (0.00%)	0 (0.00%)
Intermediate (9-14y)	7 (50.00%)	5 (62.50%)	2 (33.33%)
High (15y+)	7 (50.00%)	3 (37.50%)	4 (66.67%)
Right-handed n (%)	22 (100.00%)	11 (100.00%)	11 (100.00%)
*Median (IQR) AOO	16 (10)	16 (10)	N/A
Comorbidity (n)	1	1	N/A
MDD (n)	0	0	
Anxiety (n)	1	1	
*Medication use n (%)	1 (5.56%)	1 (14.29%)	N/A
*Mean (sd) LSAS	57.79 (42.56)	86.13 (27.31)	20 (25.84)

Note. AOO, age of onset; LSAS, Liebowitz Social Anxiety Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old. *Education and LSAS were not recorded in all SAD (n = 3) and HC (n = 5) participants. AOO and Medication use was unknown in 4 SAD participants.

Table 3.2.6.
Participant Demographic and Clinical Information: Münster

	All participants	**SAD (*n*=17)**	**HC (*n*=21)**
Female n (%)	13 (34.2%)	7 (41.2%)	6 (28.6%)

Mean (sd) age	32.66 (10.71)	34.76 (13.02)	30.95 (8.35)
Right-handed n (%)	38 (100%)	17 (100%)	21 (100%)
Comorbidity (n)	9	9	N/A
MDD (n)	6	6	
Anxiety (n)	5	5	
Medication use n (%)	0 (0%)	0 (0%)	N/A
*Mean (sd) LSAS	41.21 (27.39)	66.38 (15.34)	18.83 (11.05)
*Mean (sd) BDI	8.94 (8.42)	13.88 (9.34)	4.56 (4.15)

Note. BDI, Beck Depression Inventory; LSAS, Liebowitz Social Anxiety Scale; MADRS, MDD, major depressive disorder; n, number; sd, standard deviation. MDD and Anxiety don't add up to total comorbidity due to participants having both comorbidities. *BDI and LSAS were not recorded for 1 SAD participant and 3 HCs.

Table 3.2.7.
Participant Demographic and Clinical Information: Umea

	All participants	**SAD (*n*=23)**	**HC (*n*=19)**
Female n (%)	33 (78.60%)	20 (87.00%)	7 (63.60%)
Mean (sd) age	32.24 (10.24)	31.96 (9.56)	32.58 (11.26)
Education n (%)			
Low (5-8y)	1 (2.40%)	0 (0.00%)	1 (5.30%)
Intermediate (9-14y)	9 (21.40%)	4 (17.40%)	5 (26.30%)
High (15y+)	32 (76.20%)	19 (82.60%)	13 (68.40%)
Right-handed n (%)	42 (100.00%)	23 (100.00%)	19 (100.00%)
Median (IQR) AOO	15 (7)	15 (7)	N/A
Comorbidity (n)	0	0	N/A
Medication use n (%)	12 (28.57%)	12 (52.17%)	N/A
Mean (sd) LSAS	50.86 (32.53)	77.35 (15.92)	18.79 (10.89)
Mean (sd) MADRS	11.67 (7.28)	15.78 (6.61)	6.68 (4.42)

Note. AOO, age of onset; LSAS, Liebowitz Social Anxiety Scale; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old

Table 3.2.8.
Participant Demographic and Clinical Information: University of Illinois

	All participants	**SAD (*n*=11)**	**HC (*n*=11)**
Female n (%)	15 (68.2%)	8 (72.8%)	7 (63.6%)
Mean (sd) age	30.18 (9.64)	27.64 (7.45)	32.73 (11.20)
Education n (%)			
Low (5-8y)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Intermediate (9-14y)	7 (31.8%)	2 (18.2%)	5 (45.5%)
High (15y+)	15 (68.2%)	9 (81.9%)	6 (54.5%)
Right-handed n (%)	20 (90.9%)	10 (90.9%)	10 (90.9%)
Median (IQR) AOO	14 (13)	14 (13)	N/A
Comorbidity (n)	6	6	N/A
MDD (n)	1	1	
Anxiety (n)	5	5	
Medication use n (%)	0 (0.0%)	0 (0.0%)	N/A
Mean (sd) LSAS	45.14 (38.52)	80.27 (17.47)	10 (9.76)
Mean (sd) BDI	8.32 (10.33)	15.55 (10.32)	1.09 (1.64)

Note. AOO, age of onset; BDI, Beck Depression Inventory; LSAS, Liebowitz Social Anxiety Scale; MDD, major depressive disorder; n, number; sd, standard deviation; yo, years old

Table 3.2.9
Participant Demographic and Clinical Information: Uni of Chicago

	All participants	**SAD (*n*=22)**	**HC (*n*=16)**
Female n (%)	19 (50.0%)	9 (40.9%)	10 (62.5%)
Mean (sd) age	27.20 (6.91)	27.35 (7.48)	26.98 (6.27)
Education n (%)			
Low (5-8y)	4 (10.5%)	0 (0.0%)	4 (25.0%)
	12 (36.1%)	8 (36.4%)	4 (25.0%)

Intermediate (9-14y)	22 (57.9%)	14 (63.6%)	8 (50.0%)
High (15y+)			
Right-handed n (%)	38 (100.0%)	22 (100.0%)	16 (100.0%)
Comorbidity (n)	5	5	N/A
MDD (n)	1	1	
Anxiety (n)	4	4	
Medication use n (%)	0 (0.0%)	0 (0.0%)	N/A
Mean (sd) LSAS	46.53 (36.68)	75 (17.82)	7.38 (7.12)
Mean (sd) BDI	6.32 (6.31)	9.68 (6.14)	1.69 (2.44)

Note. FQ, Fear Questionnaire; IDS, Inventory of Depressive Symptomology; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard

Table 3.2.10.
Participant Demographic and Clinical Information: Uni of Michigan

	All participants	**SAD (*n*=31)**	**HC (*n*=41)**
Female n (%)	41 (56.9%)	20 (64.5%)	21 (51.2%)
Mean (sd) age	25.86 (7.23)	27.45 (7.72)	24.66 (6.68)
Education n (%)			
Low (5-8y)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Intermediate (9-14y)	25 (34.7%)	11 (35.5%)	14 (34.1%)
High (15y+)	47 (65.3%)	20 (64.5%)	27 (65.9%)
Right-handed n (%)	71 (98.6%)	31 (100.0%)	40 (97.6%)
Comorbidity (n)	7	7	N/A
MDD (n)	0	0	
Anxiety (n)	7	7	
Medication use n (%)	0 (0.0%)	0 (0.0%)	N/A
Mean (sd) LSAS	37.51 (35.47)	75.74 (15.54)	8.61 (7.98)
Mean (sd) BDI	5.81 (8.21)	12.61 (8.60)	.66 (1.04)

Note. FQ, Fear Questionnaire; IDS, Inventory of Depressive Symptomology; MADRS, Montgomery-Asberg Depression Rating Scale; MDD, major depressive disorder; n, number; sd, standard

4. Model Fit Statistics per Research Site, Cole Model and Kaufmann Model

Table 4.1.
Pearson Correlation and R^2 (Std. Error of the Estimate): Brain Age and Age, per Site

		Cole whole sample	**Cole SAD**	**Cole HC**	**Kaufmann whole sample**	**Kaufmann SAD**	**Kaufmann HC**
NESDA Amsterdam	Pearson	.826**	.959**	.818**	.466*	.320 (*p* = .600)	.480*
	R^2	.682 (6.94)	.920 (3.86)	.670 (7.24)	.217 (2.37)	.102 (3.99)	.231 (2.08)
NESDA Leiden	Pearson	.818**	.953**	.775**	.269 (*p* = .158)	.771 (*p* = .073)	.260 (*p* = .231)
	R^2	.670 (6.73)	.907 (4.81)	.601 (7.01)	.072 (2.55)	.594 (0.97)	.023 (2.79)
NESDA Groningen	Pearson	.830**	.785*	.837**	.555*	.848*	.372 (*p* = .289)
	R^2	.688 (6.45)	.616 (7.31)	.701 (6.63)	.308 (1.79)	.664 (1.44)	.139 (1.86)
Henk Leiden	Pearson	.790**	.837**	.764**	.668**	.700**	.650**
	R^2	.624 (4.53)	.700 (3.74)	.584 (5.37)	.446 (3.55)	.490 (3.45)	.423 (3.46)
CUBIC	Pearson	.703**	.825**	.595 (*p* = .053)	.219 (*p* = .328)	.194 (*p* = .568)	.304 (*p* = .363)
	R^2	.494 (5.19)	.680 (4.69)	.354 (5.37)	.048 (3.65)	.038 (2.72)	.092 (4.12)

Münster	Pearson	.903**	.946**	.815**	.573**	.526*	.684**
	R^2	.815 (5.15)	.895 (4.75)	.664 (5.65)	.328 (3.13)	.276 (3.08)	.468 (3.02)
UMEA	Pearson	.830**	.808**	.881**	.585**	.510*	.651**
	R^2	.689 (4.3)	.653 (4.9)	.776 (3.44)	.342 (3.56)	.260 (2.94)	.424 (4.04)
Uni of Illinois	Pearson	.847**	.890**	.846**	.465*	.415 (p = .204)	.620*
	R^2	.718 (4.88)	.792 (4.15)	.716 (5.38)	.216 (2.45)	.172 (3.04)	.384 (1.88)
Uni of Chicago	Pearson	.845**	.875**	.824**			
	R^2	.715 (4.36)	.765 (4.54)	.679 (3.60)			
Uni of Michigan	Pearson	.731**	.796**	.701**			
	R^2	.534 (4.75)	.634 (4.01)	.491 (5.24)			

Note. ** . Correlation is significant at the .01 level. * . Correlation is significant at the .05 level.

Table 4.2.
Mean (SD) and MAE (SD) BAG per Site

		Cole whole sample	**Cole SAD**	**Cole HC**	**Kaufmann whole sample**	**Kaufmann SAD**	**Kaufmann HC**
NESDA Amsterdam	Mean	1.13 (6.84)	4.45 (4.82)	.29 (7.10)	8.60 (8.68)	12.19 (6.93)	7.70 (8.99)
	MAE	5.32 (4.34)	5.12 (3.95)	5.37 (4.51)	10.20 (6.63)	12.19 (6.93)	9.70 (6.65)
NESDA Leiden	Mean	3.25 (6.64)	5.69 (4.90)	2.53 (6.99)	6.75 (8.90)	10.79 (9.64)	5.69 (8.60)
	MAE	5.47 (4.90)	5.69 (4.91)	5.41 (5.00)	8.79 (6.81)	11.08 (9.25)	8.19 (6.15)
NESDA Groningen	Mean	-0.22 (6.37)	0.59 (6.74)	-0.74 (6.39)	6.94 (9.72)	10.4 (7.42)	4.52 (10.74)
	MAE	5.43 (3.06)	5.34 (3.55)	5.49 (2.89)	10.31 (5.72)	10.4 (7.42)	10.24 (4.63)
Henk Leiden	Mean	-1.72 (4.93)	-2.47 (4.31)	-0.93 (5.52)	11.79 (5.92)	12.51 (5.81)	11.08 (6.10)
	MAE	3.96 (3.35)	3.85 (3.06)	4.09 (3.70)	11.85 (5.80)	12.51 (5.81)	11.19 (5.87)
CUBIC	Mean	1.45 (5.57)	3.06 (4.62)	-0.15 (6.18)	14.04 (7.44)	15.66 (7.69)	12.42 (7.18)
	MAE	4.47 (3.51)	4.66 (2.80)	4.29 (4.24)	14.04 (7.44)	15.66 (7.69)	12.42 (7.18)
Münster	Mean	2.06 (5.08)	2.05 (4.62)	2.06 (5.54)	10.06 (9.37)	8.21 (11.57)	11.55 (7.07)
	MAE	4.1 (3.59)	3.9 (3.10)	4.27 (4.00)	11.94 (6.73)	12.02 (7.21)	11.87 (6.50)
UMEA	Mean	-7.49 (5.78)	-7.91 (5.65)	-6.97 (6.05)	12.47 (8.26)	13.09 (8.34)	11.81 (8.3)
	MAE	8.05 (4.94)	8.50 (4.67)	7.51 (5.33)	13.09 (7.20)	13.85 (6.94)	12.29 (7.53)
Uni of Illinois	Mean	-1.53 (5.18)	0.00 (3.95)	-3.06 (5.97)	14.29 (8.72)	16.88 (6.78)	11.7 (9.95)
	MAE	4.42 (2.97)	3.3 (1.89)	5.54 (3.49)	14.79 (7.79)	16.88 (6.78)	12.71 (8.49)
Uni of Chicago	Mean	0.21 (4.3)	1.33 (4.45)	-1.31 (3.68)			

	MAE	3.27 (2.75)	3.73 (2.66)	2.64 (2.82)
Uni of Michigan	Mean	0.4 (5.20)	-0.85 (4.68)	1.34 (5.42)
	MAE	3.98 (3.33)	3.96 (2.55)	3.99 (3.85)

5. Assumptions

5.1. Normality

The below tables demonstrate skewness and kurtosis for BAG, age, mean centred age^2, age of onset and LSAS score for the Cole model and the Kaufmann model. No severe skewness and kurtosis issues were observed.

Table 5.1.1.
Assessment of Normality Using Skewness and Kurtosis (Whole Sample)

Variable	**Cole skewness**	**Cole kurtosis**	**Kaufmann skewness**	**Kaufmann kurtosis**
Brain age gap	0.076 (0.13)	1.096 (0.26)	-0.503 (0.16)	0.042 (0.31)
Age	0.769 (0.13)	-0.319 (0.26)	0.514 (0.16)	-0.674 (0.31)
Mean centred age squared	2.779 (0.13)	10.533 (.26)	2.442 (0.16)	8.877 (0.31)
LSAS (SAD only)	0.168 (0.21)	0.157 (0.42)	1.110 (0.28)	1.914 (0.55)

Note. LSAS; Liebowitz Social Anxiety Scale.

The below normal Q-Q plots show that the assumption of normality has been upheld.

Figure 5.1.1.
Normal Q-Q Plots for BAG in the Full Sample
A – Cole Model B – Kaufmann Model

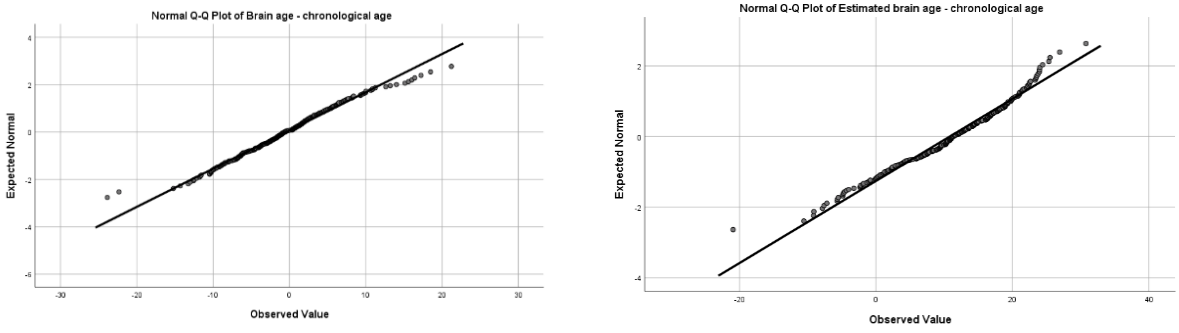


Figure 5.1.2.
Normal Q-Q Plots for Age in the Full Sample

A – Cole Model

B – Kaufmann Model

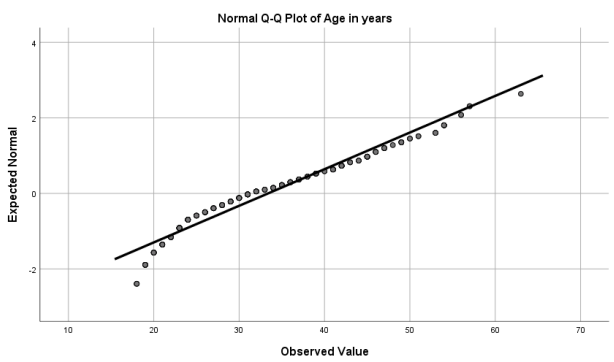
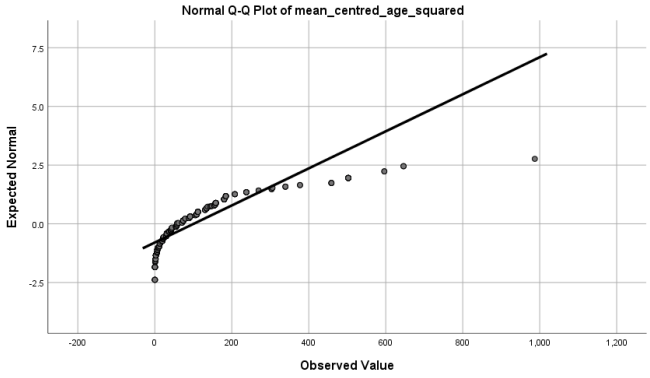


Figure 5.1.3.
Normal Q-Q Plots for Mean Centred Age^2 in the Full Sample

A – Cole model



B – Kaufmann model

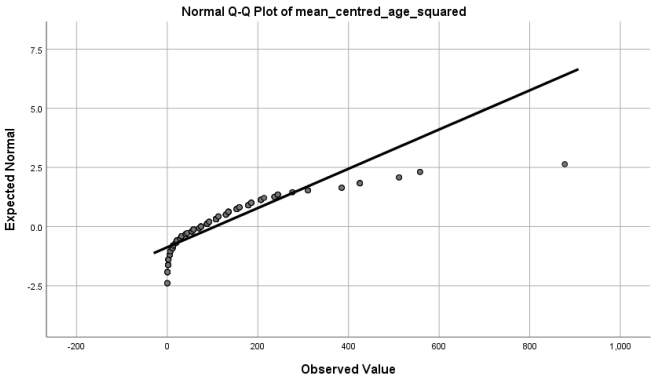
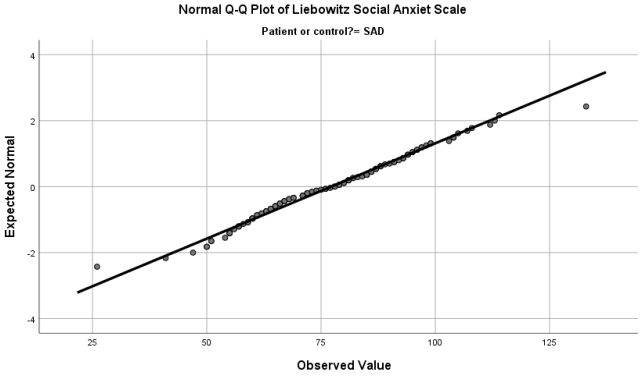
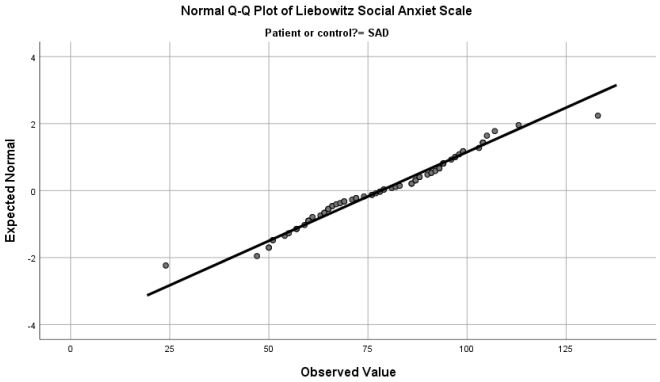


Figure 5.1.4.
Normal Q-Q Plots for LSAS in SAD Patients

A – Cole model



B – Kaufmann model



The below boxplots demonstrate that there are no problematic outliers.

Figure 5.1.5.
Boxplot Showing BAG in Patients and HCs

A – Cole model

B – Kaufmann model

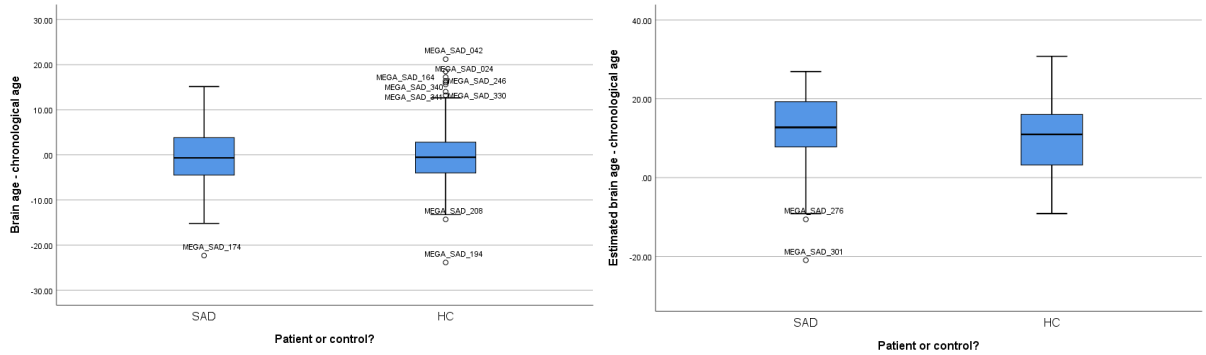


Figure 5.1.6.
Boxplot Showing Age in Patients and HCs

A – Cole model

B – Kaufmann model

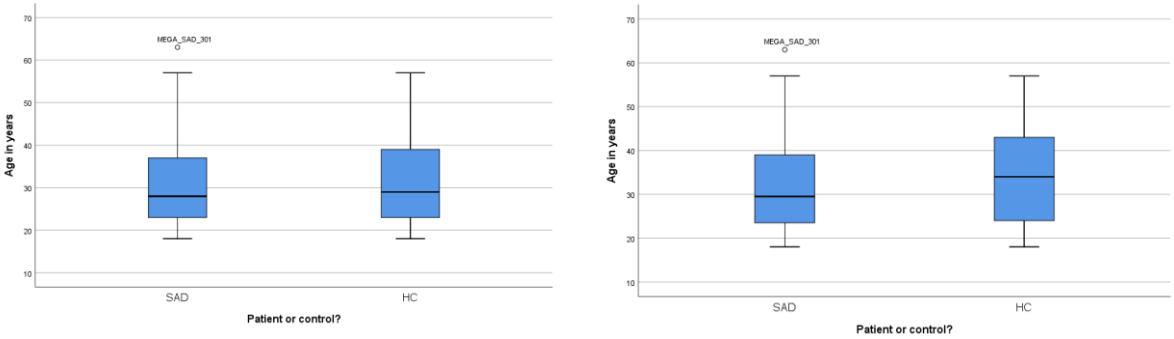


Figure 5.1.7.
Boxplot Showing Mean Centred Age^2 in Patients and HCs

A – Cole model

B – Kaufmann model

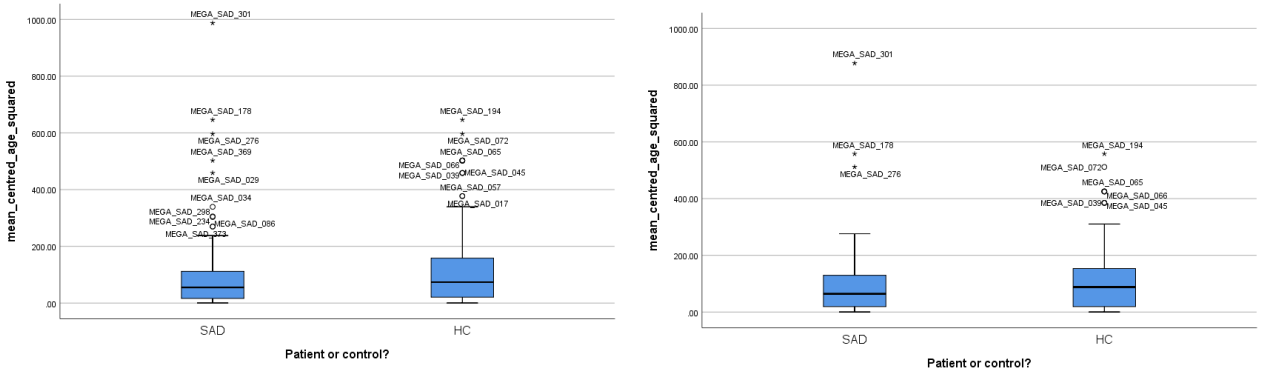
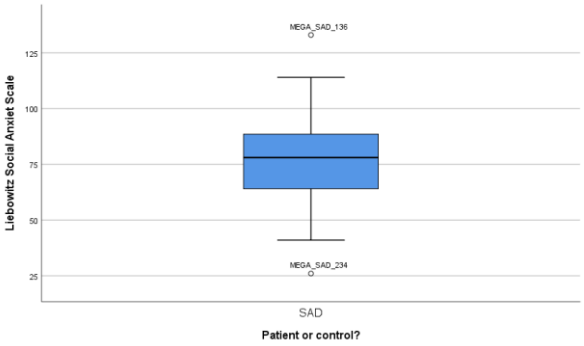
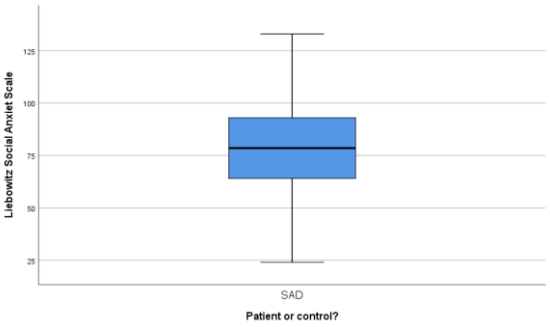


Figure 5.1.8.
Boxplot Showing LSAS in SAD Patients and HCs

A – Cole model



B – Kaufmann model

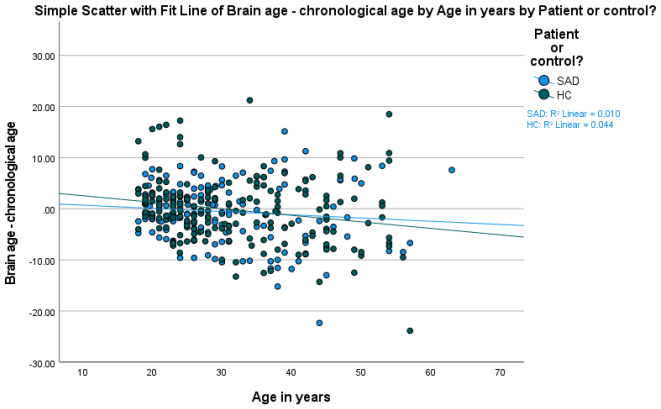


5.2. Homogeneity of Regression Slopes:

The assumption of homogeneity of regression slopes was assessed in a variety of variables using scatterplots. The assumption has been upheld, as the slopes appear to be very similar between the groups (SAD and HC) and there does not appear to be any major violations. Of note, the Cole model scatterplot demonstrates a lack of age bias when predicting brain age. A diagnosis-by-age interaction was included in the regression models to account for potential heterogeneity in regression slopes.

Figure 5.2.1.
Simple Scatterplots Showing Age in Years in SAD Patients and HCs

A – Cole model



B – Kaufmann model

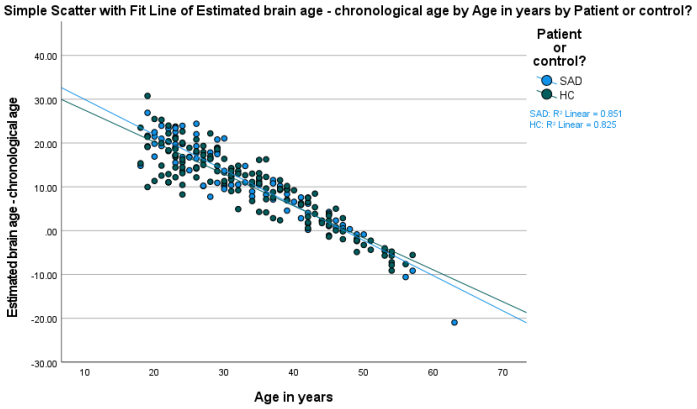
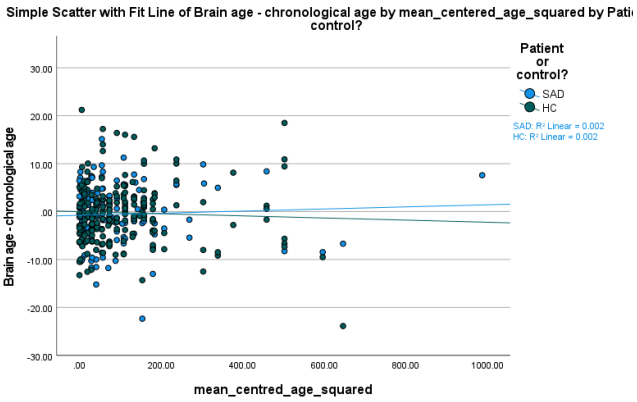


Figure 5.2.2.
Simple Scatterplots Showing Mean Centred Age^2 in SAD Patients and HCs

A – Cole model



B – Kaufmann model

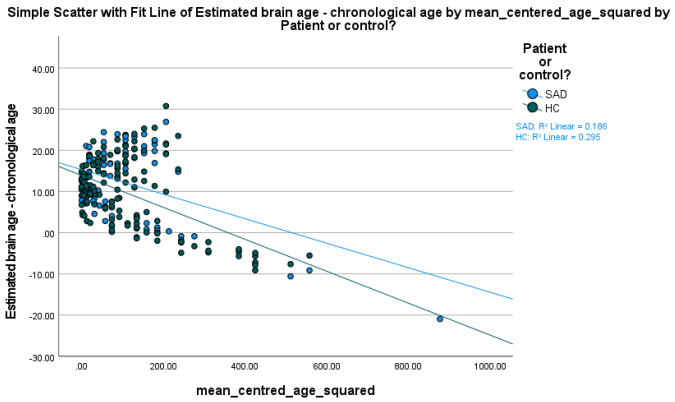
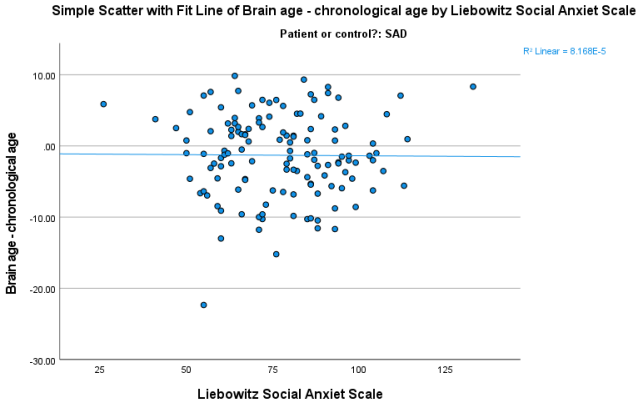
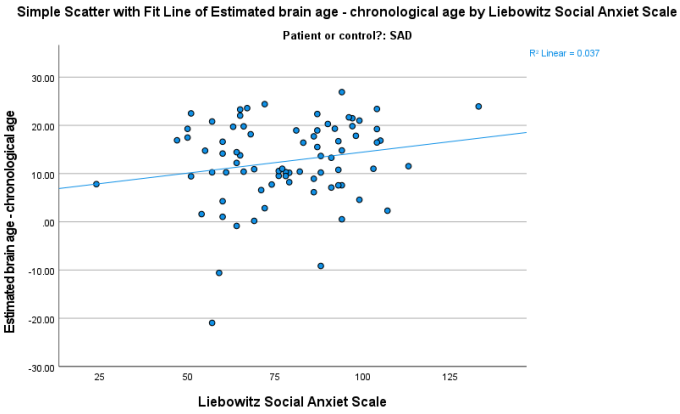


Figure 5.2.3.
Simple Scatterplots Showing LSAS in SAD Patients

A – Cole model



B – Kaufmann model



5.3. Homogeneity of Variances

5.3.1. Homogeneity of Variances, Research Question 1, Hypothesis 1b (The SAD Group will Have a Significantly Larger BAG than the HC Group)

Levene's test of equality of error variances indicated that the between-group variances were homogenous in the Cole model (F = 1.059, p = .393) and the Kaufmann model (F = 1.697, p = .053), therefore upholding the assumption of homogeneity of variance.

5.3.2. Homogeneity of Variances (Age, Mean Centred Age^2, Sex)

Levene's test of equality of error variances was not significant for all covariates for both models (p > .05), except for age in the Cole model (p = .011). Moreover, all the covariate ANOVAs are not significant, except for age in the Kaufmann model (p = .033). Therefore, the assumption of homogeneity of variance has been upheld.

5.3.3. Homogeneity of Variances, Sub-analysis 1: Psychiatric Comorbidities

Levene's test of equality of error variances indicated that the between-group variances were homogenous for the Cole model (F = 0.865, p = .659), therefore upholding the assumption of homogeneity of variance. However, this was not the case for the Kaufmann model, F = 1.740, p = .033.

Table 5.3.1.
Levene's Test of Equality of Error Variances for Sub-Analysis 1, Kaufmann Model: Psychiatric Comorbidities

F	df1	df2	*p*
1.740	19	194	.033

Note. Dependent variable: brain age gap.

Sub-analysis 1 was rerun in the Kaufmann database, using a leave one research site out approach, to see if the above Levene's test violation could be rectified. When the analysis was rerun without the University of Münster site the Levene's assumption was no longer violated (p = .083) and the analysis outcome resembled those observed in sub-analysis 1 in the results section of the dissertation (SAD group with comorbidities: n=49, β = 1.506, p = .157, η_p^2 = .012; SAD group without comorbidity: n=10, β = 1.555, p = .004, η_p^2 = .048).

SAD patients without comorbidities had a significantly higher BAG than HCs (*n*=117) by one and a half years (mean difference = 1.555, *SE* = 0.54, *p* = .013).

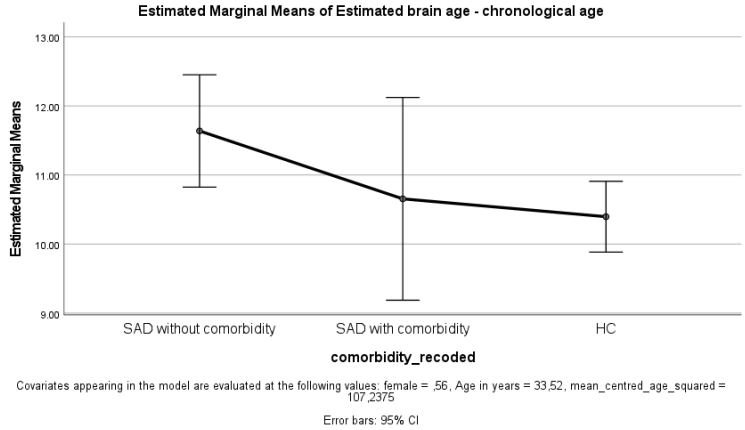
5.3.4. Homogeneity of Variances, Sub-analysis 2: Symptom Severity and Medication Use

Levene's test of equality of error variances indicated that the between-group variances were homogenous for the Cole model (F = 1.115, *p* = .354) and the Kaufmann model (F = 1.592, *p* = .136), therefore upholding the assumption of homogeneity of variance.

6. Sub-analysis 1, Kaufmann Model

Figure 6.1.

Estimated Marginal Means of BAG in SAD Without Comorbidity, SAD With Comorbidity and HCs, Kaufmann Model



7. Post Hoc Analysis

The two main analyses investigating a) whether there is a positive BAG in the SAD group, and b) whether the SAD group has a significantly larger BAG than the HC group, were rerun in the Cole database without two research sites (University of Chicago and University of Michigan) that had been excluded during QC in the Kaufmann database.

The one-sample t-test showed that the SAD group (*n*=102, research site *n*=8) did not have a positive BAG (*M* = -0.90, *SD* = 6.52) that was significantly different from 0, *t*(101) = -1.398, *p* = .165. Similarly, the HC group (*n*=140, research site *n*=8) did not have a positive

BAG (*M* = -0.59, *SD* = 6.84) that was different from 0, *t*(139) = -1.019, *p* = .310. In addition, SAD did not demonstrate a significantly larger BAG than the HC group, β = 0.815, *p* = .283. Research site showed a higher η_p^2 than in the main analysis including the two research sites (η_p^2= 0.321). Of note, the SAD and HC EM means were smaller (SAD EM mean = 0.137, *SE* = 0.58; HC EM mean = -0.677, *SE* = 0.49) than in the main analysis, suggesting that there is a large variation between research scan sites in terms of brain age.