THE ROLE OF BASAL CORTISOL ELEVATIONS IN ALZHEIMER’S DISEASE:
PRESENCE AND PREDICTION OF ASSOCIATED PATHOLOGY

A Dissertation
Presented To
The Academic Faculty

By

Ursula G. Saelzler

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Psychology

Georgia Institute of Technology
December, 2020

Copyright © Ursula Saelzler 2020
The Role of Basal Cortisol Elevations in Alzheimer’s Disease: Presence and Prediction of Associated Pathology

Approved by:

Dr. Scott Moffat, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Paul Verhaeghen
School of Psychology
Georgia Institute of Technology

Date Approved: July 29, 2020
ACKNOWLEDGEMENTS

Graduate school and the creation of this document has been quite the journey and I would like to take a moment to thank the individuals that helped me through this process. First, I would like to thank my dissertation committee. Dr. Paul Verhaeghen for his attention to detail, guidance on how to conduct a meta-analysis, and unwavering positive support that I was always grateful for. Dr. Thackery Brown for taking the time to guide me through an entirely new statistical approach and continuing to support me when I changed the analyses entirely. Dr. Mark Wheeler for making sure that I kept the basics in mind and did not lose the forest for the trees. Dr. Matthew Panizzon for his thorough feedback on each document and challenging me to be clear, concise, and consistent in my writing and research. Dr. Scott Moffat, my advisor and dissertation chair, for his feedback on each and every novel of a draft that I sent (and, in turn, wisely convincing me that my committee would rather a shorter document), as well as his support and confidence in me through this entire process.

I would also like to thank my parents for encouraging me to pursue a higher degree and supporting me through this whole process. I would like to thank my friends who have helped me more than they will ever know. Finally, I would like to thank Brandon for sticking by my side through this process. I genuinely couldn’t have done this without you.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................ iii

LIST OF TABLES .................................................................................................................. viii

LIST OF FIGURES ............................................................................................................... ix

LIST OF ABBREVIATIONS ................................................................................................. xi

SUMMARY .......................................................................................................................... xiii

CHAPTER 1: INTRODUCTION ............................................................................................... 1

Defining AD ......................................................................................................................... 1

Amyloid Cascade Hypothesis ............................................................................................. 3

Amyloid production and Aggregation ................................................................................ 4

In Vivo Imaging ...................................................................................................................... 7

Validation of PIB .................................................................................................................. 8

Quantification of PIB .......................................................................................................... 8

Relationship with Cognitive Status .................................................................................... 9

Relationship with Cognitive Decline .................................................................................. 10

Prodromal and Preclinical Stages ....................................................................................... 11

Cortisol as an AD Biomarker .............................................................................................. 12

Relationship with Cognitive Status .................................................................................... 12

Cortisol and Amyloid Interaction ...................................................................................... 14
In Vitro ........................................................................................................................................... 14
Rodent Literature ................................................................................................................................... 14
Depression........................................................................................................................................... 15
Direct Measurement of Cortisol........................................................................................................... 16
Current Studies...................................................................................................................................... 17

CHAPTER 2: CORTISOL IN THE ALZHEIMER CLINICAL SYNDROME .......... 19

Method .............................................................................................................................................. 23
Literature Search ................................................................................................................................... 23
Article Recording and Coding ............................................................................................................... 30
Addition of the BLSA Data ................................................................................................................... 38
Analytic Approach ............................................................................................................................... 39

Results................................................................................................................................................... 43
Analysis One: Basal Cortisol Differences between AD and NCC............................... 43
Analysis Two: Elimination of Influential Effects ................................................................. 46
Analysis Three: Effect of Fluid Type .......................................................................................... 48
Analysis Four: Effect of Sample Demographics ................................................................. 50
Analysis Five: Effect of Dementia Severity ............................................................................... 52

Discussion .......................................................................................................................................... 54
Unexamined Influences ....................................................................................................................... 67

Limitations ........................................................................................................................................... 70
Summary and Conclusions ........................................................................................................ 71

CHAPTER 3: CORTISOL AND AMYLOID DEPOSITION RISK ........................................ 72

Prediction of Aβ Accumulation ............................................................................................ 72

Current Investigation ............................................................................................................ 73

Method ................................................................................................................................... 75

Analytic Approach ................................................................................................................ 81

Results .................................................................................................................................. 83

Estimation of Event Time ....................................................................................................... 85

Determination of Eligible UCR Measures ........................................................................... 87

Cortisol Summary Measures ............................................................................................... 87

Cox Regression Models ......................................................................................................... 90

Post-hoc Analyses ................................................................................................................ 94

Discussion .............................................................................................................................. 96

ALH Framework .................................................................................................................. 96

Analytic Decisions and Assumptions .................................................................................. 101

Strengths ............................................................................................................................... 105

Future Directions ................................................................................................................ 106

Conclusions ........................................................................................................................... 107

CHAPTER 4: GENERAL DISCUSSION .............................................................................. 108

Alternative Interpretations ................................................................................................. 109
Directionality of Effects.............................................................................................................. 109

Uniformity of effects.................................................................................................................. 111

Etiology of Cortisol Elevations.................................................................................................. 112

Methodological Considerations ............................................................................................... 113

Unidimensionality of Basal Cortisol....................................................................................... 113

PIB and Revisions to theAmyloid Cascade Hypothesis ......................................................... 114

Cautionary Notes ..................................................................................................................... 115

Conclusions ............................................................................................................................... 116

REFERENCES ........................................................................................................................... 117
LIST OF TABLES

Table 1: Analysis Elements Vocabulary ................................................................. 25
Table 2: Included Effects and Covariates by Article .............................................. 35
Table 3: Meta-Analytic Summary Results ............................................................... 44
Table 4: Specification of Terms .............................................................................. 80
Table 5: Demographic Comparisons ..................................................................... 84
Table 6: Cox Regression Analysis Results .............................................................. 93
Table 7: Post-hoc Cox Regression Analysis Results ............................................... 95
LIST OF FIGURES

Figure 1. The amyloidogenic pathway of APP metabolism.................................................. 6
Figure 2. The development of Aβ fibrils.............................................................................. 6
Figure 3. Process for sorting articles for inclusion in the present meta-analysis............... 29
Figure 4. The ratio of AD and NCC participants used for effect size calculation.......... 40
Figure 5. The distribution of the primary study effect sizes.............................................. 40
Figure 6. The relationship between sample pair size (N) and the variance of the effect
size. ........................................................................................................................................... 41
Figure 7. Scatterplot of the relationship between the SMD of cortisol levels for the AD
and NCC samples as a function of year of publication...................................................... 41
Figure 8. The influence diagnostics from Analysis One...................................................... 45
Figure 9. The funnel plot for the effects included in Analyses Two through Four.......... 47
Figure 10. Forest plot of the effects included in Analyses Two and Analysis Three....... 49
Figure 11. The relationships between the weighted mean covariates of the sample pairs
and the SMD of cortisol levels between the AD and NCC samples................................. 51
Figure 12. Scatterplot showing the relationship between the average MMSE score for the
AD sample and the SMD of cortisol for the AD and NCC samples................................. 53
Figure 13. Illustration of three hypothetical relationships between cortisol and age for
adults with AD and NCCs....................................................................................................... 63
Figure 14. The count, by year and type of collection, of the PIB and UCR collections
only for those participants (N = 145) with at least one UCR measure and one PIB
measure. ................................................................................................................................... 78
Figure 15. The timing of UCR and PIB measurements in relation to the baseline PIB measurement for all participants with at least one UCR measure and at least one PIB measure.

Figure 16. Plot of PIB measurements for participants with at least one PIB measurement and one UCR measurement plotted by the number of years away from PIB baseline.

Figure 17. Timing of UCR measurement and event time (rounded to the nearest year).
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>Alzheimer's Association</td>
</tr>
<tr>
<td>Aβ</td>
<td>Amyloid-beta</td>
</tr>
<tr>
<td>ACTH</td>
<td>Adrenocorticorophic hormone</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer's disease</td>
</tr>
<tr>
<td>ADNI</td>
<td>Alzheimer’s Disease Neuroimaging Initiative</td>
</tr>
<tr>
<td>ADRDA</td>
<td>Alzheimer’s Disease and Related Disorders Association</td>
</tr>
<tr>
<td>AIBL</td>
<td>Australian Imaging, Biomarkers and Lifestyle</td>
</tr>
<tr>
<td>ALH</td>
<td>Allostatic Load Hypothesis</td>
</tr>
<tr>
<td>APOE</td>
<td>Apolipoprotein E</td>
</tr>
<tr>
<td>APP</td>
<td>Amyloid precursor protein</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>BL</td>
<td>Baseline</td>
</tr>
<tr>
<td>BLSA</td>
<td>Baltimore Longitudinal Study of Aging</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>CAR</td>
<td>Cortisol Awakening Response</td>
</tr>
<tr>
<td>CESD</td>
<td>Center for Epidemiological Studies Depression Scale</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CSF</td>
<td>Cerebrospinal fluid</td>
</tr>
<tr>
<td>DSM</td>
<td>Diagnostic and Statistical Manual of Mental Disorders</td>
</tr>
<tr>
<td>DVR</td>
<td>Distribution Volume Ratio</td>
</tr>
<tr>
<td>cDVR</td>
<td>Cortical Distribution Volume Ratio</td>
</tr>
<tr>
<td>GCH</td>
<td>Glucocorticoid Cascade Hypothesis</td>
</tr>
<tr>
<td>HDL</td>
<td>High density lipoprotein</td>
</tr>
<tr>
<td>HPA</td>
<td>Hypothalamic pituitary adrenal axis</td>
</tr>
<tr>
<td>HPLC-MS</td>
<td>High-performance liquid chromatography-mass spectrometry</td>
</tr>
<tr>
<td>HR</td>
<td>Hazard Ratio</td>
</tr>
<tr>
<td>HTN</td>
<td>Hypertension</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Diseases</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>LDL</td>
<td>Low density lipoprotein</td>
</tr>
<tr>
<td>MCI</td>
<td>Mild Cognitive Impairment</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini Mental Status Examination</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>NCC</td>
<td>Normal cognition control</td>
</tr>
<tr>
<td>NIA</td>
<td>National Institute on Aging</td>
</tr>
<tr>
<td>NINCDS</td>
<td>National Institute of Neurological and Communicative Disorders and Stroke</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>PIB</td>
<td>Pittsburgh Imaging Compound-B</td>
</tr>
<tr>
<td>RIA</td>
<td>Radioimmunoassay</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard error of the mean</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>SMD</td>
<td>Standardized mean difference</td>
</tr>
<tr>
<td>SUVR</td>
<td>Standardized uptake value ratio</td>
</tr>
<tr>
<td>UCR</td>
<td>Urinary free cortisol to creatinine ratio</td>
</tr>
<tr>
<td>UFC</td>
<td>Urinary free cortisol</td>
</tr>
</tbody>
</table>
Elevations in the glucocorticoid steroid hormone cortisol have long been suspected to be characteristic of Alzheimer’s disease, and there is a renewed interest in the contribution of cortisol to the progression of the disease. However, to date, no systematic examination of basal cortisol levels in AD has been conducted. Further, although cortisol elevations are hypothesized to facilitate cerebral amyloid-beta accumulation, direct examinations of this relationship have been restricted to samples already showing evidence of amyloid accumulation. The two studies included in this dissertation directly addressed these issues. First, a comprehensive meta-analysis was conducted to compare basal cortisol levels between cognitively normal and demented older adults. The cortisol levels in the demented samples were significantly higher than those of the cognitively normal samples ($d = 0.39, Q_M = 34.73, p < 0.001$) and this difference was not moderated by the age, sex, or cognitive performance of the included samples. The second study applied a Cox regression analysis to data collected from the Baltimore Longitudinal Study of Aging to test whether basal cortisol measures collected prior to imaging evidence of amyloid-beta deposition would predict the risk of subsequent amyloid-beta accumulation. Both the cortisol slope and the cortisol level measured nearest to the estimated event time were significantly positively associated with the risk of amyloid-beta accumulation (HR = 1.31, $p = 0.04$ and HR = 1.03, $p = 0.01$ respectively). Taken together these studies suggest that basal cortisol elevations are often observed in Alzheimer’s disease and that the influence of cortisol on disease progression may begin even before detection by neuroimaging markers.
CHAPTER 1: INTRODUCTION

Alzheimer’s disease (AD) is a neurodegenerative disease affecting an estimated 5.8 million Americans (Alzheimer's Association, 2019) with projected increases to 13.8 million individuals by 2050 (Hebert et al., 2013). Of those currently affected, 5.6 million are over the age of 65 with the prevalence of the disease increasing with age, such that 32% of people over age 85 are impacted (Hebert et al., 2013). Given the rapid increase in the percentage of the population over 65, primarily driven by the aging of the baby boomer generation (Ortman et al., 2014), the prevalence of AD amongst older adults poses a major public health concern (Alzheimer's Association, 2019).

Defining AD

The clinical and histological aspects of AD, including progressive memory impairment and the accumulation of plaques were first described by Alois Alzheimer in 1906 (Maurer et al., 1997). At present, it is understood that these clinical and biological markers are produced by one of the two types of AD: familial (also referred to as presenile or early onset) and sporadic (or late onset). Familial AD is characterized by genetic mutations on chromosomes 1, 14, and 21 and the onset of cognitive impairment occurring around age 45 (de la Torre, 2016). In contrast, sporadic or late onset AD typically manifests after age 65 and has genetic factors which increase the susceptibility to AD but do not guarantee its development (de la Torre, 2016). Sporadic AD accounts for an estimated 95% of AD cases (Wu, L. et al., 2012) and is the focus of the research proposed here.
Beginning in 1984, a number of workgroups have attempted to create a common vernacular and criteria to describe the clinical and biological characteristics of AD (e.g., Albert et al., 2011; McKhann et al., 1984; McKhann et al., 2011; Mirra et al., 1991; Morris et al., 1989; Morris et al., 1988; Sperling et al., 2011). The first workgroup was established by the NINCDS and the ADRDA with the goal of establishing and describing the clinical criteria necessary for the diagnosis of AD. This group outlined criteria for diagnoses of possible, probable and definite AD such that a diagnosis of probable AD could be made in vivo when progressive deterioration of cognitive functions was present in the absence of any other identifiable cause (e.g., Parkinson’s disease, drug intoxication, thyroid disease etc.). However, in order to make a definite AD diagnosis, histopathologic confirmation, generally at autopsy, was required (McKhann et al., 1984). The restriction of a definitive AD diagnosis to those with histopathologic confirmation set the precedence for the distinction between the clinical manifestations of AD and the biological etiology.

The most recent workgroup commissioned through the NIA-AA created a set of recommendations to serve as a framework for AD research. This group emphasized the importance of the distinction between the Alzheimer’s clinical syndrome, defined as the cluster of clinical symptoms and Alzheimer’s disease, the biological etiology of the clinical symptoms (identified through biomarkers). In order to reflect the revised distinction wherein the dementia syndrome classification is determined entirely agnostic to the presence or absence of AD biomarkers, the terminology for the three Alzheimer’s clinical syndrome stages were amended to cognitively unimpaired, mild cognitive impairment (MCI), and dementia (Jack et al., 2018).
To create a common vernacular for describing the aggregate of pathophysiological changes in AD, the 2018 group proposed the "AT(N)" system of biomarker classification. This system assigns a biomarker profile to individuals based on the binary presence or absence of AD biomarkers. Specifically, the presence (+), absence (-), or unmeasured (*) nature of: Aß plaques (identified through cortical amyloid PET ligand binding or low CSF Aß42), labeled "A"; fibrillar tau (identified through elevated CSF phosphorylated tau (P-tau) or cortical tau PET ligand binding), labeled "T"; and neurodegeneration or neuronal injury (identified through CSF T-tau, fluorodeoxyglucose, PET hypometabolism or atrophy on MRI), labeled "(N)" is indicated. The "(N)" appears in parentheses because these biomarkers are not specific to AD (Jack et al., 2018).

Individuals are assigned to one of eight possible AT(N) profiles, such as "A+T+(N)-" which specify membership in one of the three aggregate biomarker categories of: normal AD biomarkers, non-AD pathologic change, or in the Alzheimer’s continuum. The group notes an awareness of numerous limitations with the currently available biomarkers and stress the need for further examination of these biomarkers (as well as other low-cost and minimally invasive biomarkers) in population-based studies before this proposed research framework can extend into clinical practice (Jack et al., 2018).

**Amyloid Cascade Hypothesis**

A key feature of the AT(N) system is that it restricts assignment to the AD continuum to those that are A+, revealing the underlying assumption that elevated Aß accumulation is requisite to the classification of AD. The workgroup was explicit, however, that the classification system does not require, nor assert, that cerebral Aß
accumulation is the driver of AD progression (Jack et al., 2018). This distinction reflects a departure from the historically dominant approach to AD research. Specifically, the notion that the accumulation of cerebral Aβ is the driver of AD was originally articulated by Hardy and Higgins (1992) as the Amyloid Cascade Hypothesis. While several modifications have been made to this hypothesis in the intervening decades (Hardy & Selkoe, 2002; Lambert et al., 1998; Selkoe & Hardy, 2016; Walsh & Selkoe, 2007), the assumption that Aβ, in some form, drives disease progression is still at the forefront of AD research.

**Amyloid production and Aggregation**

Aβ is one product of the metabolism of the APP. The modern understanding is that APP is first cleaved by α- or β-secretase. If cleaved by α-secretase, APP is metabolized without the creation of Aβ (i.e., the non-amyloidogenic pathway). Conversely, if APP is cleaved by β-secretase the resulting amino acid fragment is then cleaved by γ-secretase at either γ-, ζ-, or ε-sites creating Aβ peptides that are 37 – 43 amino acid residues long (see Figure 1). The 40-amino acid peptide (Aβ40) is the most abundantly produced Aβ peptide (accounting for approximately 90% of Aβ). Fewer Aβ38 and Aβ42 peptides are produced, and even fewer Aβ37 and Aβ43 peptides (Steiner et al., 2018; Yuksel & Tacal, 2019). AD research focuses primarily on Aβ40 and Aβ42 due to the prevalence of both peptides in amyloid plaques coupled with the high propensity of Aβ42 to aggregate (Steiner et al., 2018).

In the classical model of Aβ fibril formation, these Aβ monomers (i.e. single 40 and 42 amino acid residue long peptides) follow a nucleation-dependent polymerization process whereby monomers aggregate to a variety of oligomeric structures (Haass &
Selkoe, 2007; Ono, 2018; Walsh & Selkoe, 2007). These oligomers serve as a nucleus which creates a template for incorporating more monomers into a highly ordered beta-sheet fibrillar structure, producing protofibrils and mature fibrils (Brännström et al., 2018; Ono, 2018) (see Figure 2).
Figure 1. The amyloidogenic pathway of APP metabolism. The full APP protein is first cleaved by $\beta$-secretase. The resulting fragment is then cleaved at one of three sites (the $\gamma$-, $\zeta$-, or $\varepsilon$-site), resulting in a monomer of A$\beta$. Image adapted from Figure 1 of Walsh and Selkoe (2007).

Figure 2. The development of A$\beta$ fibrils. Monomers of A$\beta$ aggregate to form oligomeric nuclei, which elongate to promote fibril formation. Adapted from Figure 7 from Kumar et al. (2011).
Finally, these varying A\(\beta\) structures aggregate to A\(\beta\) plaques (Abrahamson et al., 2019; Tschanz, 2011). Although in vitro studies have consistently demonstrated that A\(\beta\), in a variety of structures, is directly neurotoxic (Goodman et al., 1996), the exact mechanism of neurotoxicity in vivo is still debated.

**A\(\beta\) deposition scenarios.**

There is little dispute that A\(\beta\) accumulation is a hallmark of AD, however, there are still many questions regarding the role and timing of A\(\beta\) accumulation in the progression of AD. Researchers have framed the original Amyloid Cascade Hypothesis as the *A\(\beta\) as driver* scenario. In this scenario, continuous deposition of A\(\beta\) is required for AD progression and the speed of disease progression is directly related to the rate of deposition. (Karran et al., 2011). Two other scenarios, the *A\(\beta\)-trigger* scenario and the *A\(\beta\)-threshold* scenario, propose that the AD progression begins when A\(\beta\) deposition reaches some benchmark. From this point, the *A\(\beta\)-trigger* scenario asserts that once the progression of AD has started, progress cannot be stopped or slowed regardless of alterations in A\(\beta\) deposition. Conversely, the *A\(\beta\)-threshold* scenario asserts that even after starting, disease progression can be slowed or stopped if therapeutics successfully resolve amyloid plaques to a level below the threshold (Karran et al., 2011).

**In Vivo Imaging**

Questions regarding the timing and role of A\(\beta\) deposition in AD are not new, however, until 2004, A\(\beta\) deposition could only be measured at autopsy. At the turn of the century, scientists modified the A\(\beta\) dye thioflavin-T, creating a radiolabeled compound capable of crossing the blood brain barrier that could be imaged using PET. This compound, N-methyl-[11C]2-(4’-methylaminophenyl)-6-hydroxybenzothiazole (termed
"Pittsburgh Compound-B" or PIB), showed high affinity for synthetic aggregated Aβ and rapid clearance from rodent brain tissue (Klunk et al., 2001; Mathis et al., 2002). Several years later, the first "proof-of-concept " study in humans was conducted, demonstrating marked retention of PIB in several brain regions of AD patients compared to healthy adults in vivo (Klunk et al., 2004).

**Validation of PIB**

Although the initial proof-of-concept report of PIB binding in humans described an in vivo pattern of Aβ deposition consistent with earlier post-mortem descriptions of Aβ deposition patterns, the initial study did not follow participants to autopsy (Klunk et al., 2004). In order to further validate PIB as an in vivo marker of Aβ deposition, subsequent case studies compared individuals’ PIB scans with their Aβ levels at autopsy. These studies reported positive correlations between Aβ deposition at autopsy and regional PIB binding in patients with Lewy body dementia (Bacskai et al., 2007), and AD (Ikonomovic et al., 2008; c.f., Rosen et al., 2010). In the following decade, larger studies including participants from a broader cognitive spectrum further examined the validity of PIB (e.g., Bergeron et al., 2018; Hatsuta et al., 2015; Seo et al., 2017). These examinations reported positive, but not perfect, correlations between PIB and post-mortem Aβ, typically ranging from correlations of 0.30 to 0.90 depending on the region examined and quantification method of post-mortem Aβ.

**Quantification of PIB**

Like other PET radiotracers, PIB is frequently quantified using SUVR or DVR. SUVR and DVR can be calculated for specific regions of interest by comparing the strength of the PET signal in the specified region against the strength of the signal in a
reference region, typically averaged over a specified period of time. The cerebellum is typically used as the reference region for PIB due to the low deposition of Aβ plaques to this area (Joachim et al., 1989; Klunk et al., 2004; Yamaguchi et al., 1989), therefore higher SUVR or DVR values are indicative of greater plaque deposition.

These values are either used as continuous measures, or more commonly, participants are dichotomized as either Aβ+ or Aβ- based on measurement specific thresholds for global uptake (typically around 1.4-1.5 for SUVR and 1.2 for DVR) (Villeneuve et al., 2015). More recently, protocols for visually assessing and dichotomizing raw PIB PET images have been developed (see Bourgeat et al., 2015). Unlike SUVR and DVR dichotomization, visual assessment methods do not require structural MRI acquisition allowing for greater inclusion due to the prevalence of MRI contraindications in the older adult population. Agreement on amyloid positivity made visually by trained raters tends to be high (91.3% for AD-like amyloid positivity in the case of Mountz et al. (2015)), as is agreement between visual classification and SUVR and/or DVR classification (Mountz et al., 2015; Ng et al., 2007; Rabinovici et al., 2011).

**Relationship with Cognitive Status.**

Mirroring the results of post-mortem Aβ quantification studies (Arnold et al., 1991; Blessed et al., 1968; Muramori et al., 1998; Roth et al., 1966), the prevalence of Aβ positivity, as indicated by in vivo PET imaging, is lowest among cognitively normal persons and highest in those with AD (e.g., Duara et al., 2019). While the absolute prevalence of Aβ positivity in cognitively normal, MCI and AD patients varies due to differences in global SUVR and DVR threshold definitions, within-study cross-group comparisons consistently show an increase in the prevalence of Aβ positivity from
cognitively normal (0-47%) to MCI (37%-72%) to AD (68 – 100%) (see Table 1 of Chételat et al. (2013). Similarly, two meta-analyses of in vivo Aβ quantification reported that 24% of non-demented adults, 53% of MCI patients, and 88% of clinically diagnosed AD patients were classified as Aβ+ (Jansen et al., 2015; Ossenkoppele et al., 2015).

**Relationship with Cognitive Decline**

Given the clear increase in the prevalence of Aβ positivity among MCI and AD patients researchers have aimed to determine whether Aβ positivity is associated with poorer cognitive functioning cross-sectionally and whether Aβ positivity can serve as a biomarker to predict conversion from normal cognition to dementia. A meta-analysis of 38 studies reported that among cognitively normal older adults Aβ positivity was associated with poorer performance across five cognitive domains cross-sectionally, and decline in four domains longitudinally (Baker et al., 2017). Later reports have further supported these findings (Bos et al., 2018; Pothier et al., 2019).

Similarly, several reports demonstrate that Aβ positivity increases the risk of progression to dementia. A study of 599 initially cognitively normal individuals demonstrated that Aβ positivity more than doubled the odds (OR: 2.43) of progressing from cognitively normal to MCI or dementia during an interval of eight years (Dang et al., 2018). Wolk et al. (2018) reported that amongst 232 participants with MCI, 53.6% of those with an Aβ+ scan at baseline converted to probable AD during the 36-month follow-up while only 22.8% of those with Aβ- scans converted during this same interval. These findings are consistent a number of other studies and meta-analytic findings reporting the progression of 40-80% of MCI Aβ+ subjects to AD in similar follow-up intervals (Ma et al., 2014; Villemagne & Chételat, 2016).
**Prodromal and Preclinical Stages**

The ability to identify in vivo those individuals who are at increased risk to develop dementia has led to a shift in AD research towards understanding the preclinical and prodromal stages of AD. While the concept of a preclinical stage of AD is not new (e.g., Hubbard et al., 1990; Hulette et al., 1998), the direct study of this period was limited prior to the development of PIB. In the context of the modern understanding of AD as a continuum, the *preclinical* stage of AD is defined by the presence of biomarkers of AD pathology (e.g. Aβ positivity) in the absence of clinical symptoms. The clinical presentation of the *prodromal* stage is largely analogous with earlier definitions of MCI due to AD, such that a designation of prodromal AD requires the presence of a pattern of clinical symptoms consistent with those observed in AD but of insufficient severity to warrant a dementia diagnosis. The current designation of prodromal AD also requires concurrent biomarkers of AD pathology (Dubois, 2000; Dubois & Albert, 2004; Dubois et al., 2016).

One of the first direct longitudinal investigations of the preclinical and prodromal phases of AD suggested that the preclinical phase of AD may begin up to two decades prior to the onset of AD. Villemagne et al. (2013) examined PIB scans from a total of 200 cognitively normal, MCI, and demented older adults every 18 months for 3-6 years. Using a prospective longitudinal design, these authors concluded that the combined preclinical and prodromal stages of AD last approximately 17 years with the majority of this time spent in the preclinical phase.
Cortisol as an AD Biomarker

The purported duration of the preclinical phase of AD has shifted research towards identifying potential accelerants of Aβ deposition with a focus on easily obtained biomarkers (Jack et al., 2018). One such biomarker that has recently gained attention is cortisol (Ahmad et al., 2019). Cortisol is a lipophilic, glucocorticoid steroid hormone produced by the adrenal gland, capable of passing through the blood-brain barrier (Dunn et al., 1981; Pardridge & Mietus, 1979). Although cortisol is involved in a number of regulatory functions (e.g. Martin, 1995, p. 289), it frequently receives attention for its role in the stress response (McEwen, 2019). Stemming from early rodent research demonstrating that chronic stress or exogenous glucocorticoid exposure produces cognitive deficits mimicking those observed in aged rodents (Arbel et al., 1994; Conrad, 2010), cortisol has long been hypothesized as a driver of age-related cognitive changes (see Lupien et al., 2018 for discussion). More recently, researchers have suggested that cortisol, specifically prolonged cortisol elevation, is the mediating factor between depression and AD (Aznar & Knudsen, 2011; Canet et al., 2018; Sotiropoulos et al., 2008).

Relationship with Cognitive Status

Owing to its hypothesized role in age-related cognitive declines, cortisol has been frequently compared between AD and control participants. A number of researchers have reported elevated cortisol levels in AD patients and it is often taken for granted that AD is consistently characterized by higher endogenous cortisol levels however, this is far from a unanimous finding in the literature (see Chapter 2).
The reported elevation of cortisol in AD patients has led researchers to examine whether cortisol alone, or as part of a biomarker profile, can accurately predict those who will develop AD. Although formal clinical diagnoses were not made, (Karlamangla et al., 2005) reported that among 538 participants aged 70-79 years at baseline enrolled in the MacArthur Study of Successful Aging, higher overnight cortisol levels were associated with a higher risk of incident cognitive impairment.

Doecke et al., (2012) used data on 961 participants from the AIBL cohort to identify a biomarker profile distinguishing AD patients from controls. Cortisol, included as a set of eight biomarkers,1 in addition to age, sex, and APOE genotype, was validated using 170 individuals from the ADNI with sensitivity and specificity of 83%. In this same sample, cortisol was identified as one of a set of six fluid biomarkers which correctly distinguished between those individuals with MCI that would progress to AD within 3 years with 80% accuracy (see eTable 6 of Lehallier et al., 2016).

Similarly, in a sample of 1,025 adult participants from the BLSA with two or more measures of 24-hour urinary free cortisol, Ennis et al. (2017) reported that mean cortisol levels were a significant predictor of AD development (HR = 1.24, p = 0.01). This association remained significant after controlling for several covariates, including APOE e4 status (HR = 1.31, p = 0.005).

1 The full set of eight biomarkers included cortisol, insulinlike growth factor binding protein, pancreatic polypeptide, interleukin-17, vascular cell adhesion modelcute 1, β2 microglobulin, epidermal growth factor receptor and carcinoembryonic antigen
Cortisol and Amyloid Interaction

The presence of cortisol elevations in those at elevated risk for the transition to AD coupled with the Aβ deposition observed in AD have led researchers to speculate that cortisol elevations may contribute to Aβ accumulation (see Libro et al., 2017 for discussion).

In Vitro

Green et al. (2006) reported that the treatment of cultured mouse neuronal N2a cells with varying concentrations of glucocorticoid hormones increased the amounts of Aβ40 and Aβ42 present in a concentration and time-dependent manner. Further, this increase in Aβ was prevented or diminished by the addition of glucocorticoid receptor antagonists. Analogous findings were reported in primary astrocyte cultures (Wang, Y. et al., 2011).

Rodent Literature

Beginning in the mid-1990s, a number of transgenic mouse models of AD pathophysiology have been developed (see Table 1 of Sadigh-Eteghad et al., 2015). Many of these models demonstrate amyloid plaque deposition with age, allowing researchers to examine in vivo biological factors which may accelerate Aβ deposition. Green et al. (2006) injected one such model, 3xTg-AD mice (Oddo et al., 2003), with a glucocorticoid hormone daily for one week. Enzyme-linked immunosorbent assay quantification revealed significant increases in both whole brain Aβ40 and Aβ42 in injected rodents. Like exogenous glucocorticoid administration, chronic stress, a driver of endogenous cortisol elevations, is capable of producing neuronal Aβ increases in rodent models (Dong et al., 2004; Kang et al., 2007). For example, Jeong et al. (2006) reported
that eight months of chronic isolation stress in APPv717rCT100 mice was sufficient to significantly increase corticosterone levels as well as increase extracellular Aβ plaque deposition and intraneuronal Aβ. Similarly, Dong et al. (2004) reported that several months of isolation stress in Tg2576 mice was sufficient to significantly increase Aβ plaques (see Dong & Csernansky, 2009 for review).

**Depression**

Individuals with depression often experience chronic cortisol elevations for months to years at a time (Murri et al., 2014; Stetler & Miller, 2011) and are consistently reported to be at increased risk of dementia (Barnes et al., 2012; Cherbuin et al., 2015; Diniz et al., 2013; Ownby et al., 2006; Rasmussen et al., 2018). These findings have led researchers to suggest that cortisol elevations may be the mediating factor between depression and AD and that this mediation may be driven by increased Aβ deposition (Canet et al., 2018; Sotiropoulos et al., 2008).

In support of this hypothesis, increased Aβ deposition has been reported in cognitively normal older adults with depression or depressive symptoms (see Harrington et al., 2015 for review). Among 1,038 cognitively normal participants, Krell-Roesch et al. (2018) reported that the score of the Beck Depression Inventory-II (Beck et al., 1996) significantly linearly correlated with global PIB SUVR. Similarly, Wu, K.-Y. et al. (2014) reported significantly increased Aβ retention in the parietal cortex and precuneus among cognitively normal older adults with a history of depression.

Similar relationships between lifetime history or depressive symptoms and Aβ deposition have been reported in MCI patients (Chung et al., 2015; Krell-Roesch et al., 2019; Lavretsky et al., 2009). Further, Brendel et al. (2015) reported that of 44 MCI
patients with concurrent subsyndromal depression and high amyloid positivity, 100% converted to AD within 30 months (see Figure 3D of Brendel et al., 2015). The exacerbation of Aβ deposition through depression appears to continue through dementia diagnosis. Rapp et al. (2006) reported that patients with a history of both major depression and AD had more neuritic Aβ plaques in the hippocampus than patients with only AD at autopsy. Similarly Meynen et al. (2010) reported that amongst 43 AD patients, the Cornell score (Alexopoulos et al., 1988) obtained an average of 3 months before death was positively correlated with neuritic Aβ plaque density throughout the cortex and in particular in the temporal cortex.

**Direct Measurement of Cortisol**

The study of depressed patients provides ample indirect evidence for a relationship between chronic cortisol elevations and AD, however, very little work has directly examined the relationship between cortisol and Aβ deposition in eucortisolemic adults. Toledo et al. (2012) first reported a relationship between endogenous cortisol levels and PIB retention in a subset of 81 subjects from the ADNI cohort. Adjusting for age, clinical diagnosis (cognitively normal, MCI and AD) and APOE genotype, every log increment of morning plasma cortisol was associated with an increase of one in the SUVR PIB summary score.

More recently, Pietrzak et al. (2017) examined the cognitive trajectories of four groups composed of a total of 416 cognitively normal older adults created through dichotomization of morning plasma cortisol levels (median split) and PIB SUVR (Aβ+ and Aβ-). The authors reported a significant Plasma Cortisol x Aβ x Time interaction effect on global cognitive performance such that the high cortisol, Aβ+ individuals
demonstrated greater declines in cognitive performance over the 72-month study period. While the findings of Pietrzak et al. (2017) do not address whether cortisol elevations exacerbate Aβ deposition, the results suggest that elevations of both Aβ and cortisol can have deleterious effects on cognition potentially leading to a clinical diagnosis of MCI due to AD or Alzheimer’s syndrome.

Using a similar dichotomization approach in the ADNI cohort, Udeh-Momoh et al. (2019) reported that individuals with high cortisol and abnormal Aβ42 were at significantly increased risk for progression from a cognitively normal state to MCI or dementia, in comparison to individuals with low cortisol and normal Aβ42 (HR = 3.67, p = 0.017).

Current Studies

In response to the accumulating evidence that cortisol may play a role in the development and/or expression of dementia, this dissertation will examine the association between basal cortisol and both the Alzheimer clinical syndrome and the pathophysiology of Alzheimer’s disease. To do so, this dissertation consists of two studies investigating (a) the presence of basal cortisol elevations in the Alzheimer clinical syndrome and (b) the association between cortisol and Aβ deposition as determined from PIB PET scans.

The first study consists of a meta-analysis of the literature comparing basal cortisol levels in cognitively normal and demented patients to determine whether basal hypercortisolism is a marker of the clinical AD syndrome, and, if so, to estimate the magnitude of this effect. Further, it examined whether sample characteristics such as age, sex or clinical severity moderate this difference and provide insight into the role of cortisol in the Alzheimer clinical syndrome.
The second study consists of a survival analysis model examining whether extended basal cortisol levels or a pattern of increasing basal cortisol levels influence the risk of Aβ accumulation among previously Aβ pathology-free middle-aged and older adults. The survival analysis also examined whether participant age moderated the risk of Aβ accumulation.
CHAPTER 2: CORTISOL IN THE ALZHEIMER CLINICAL SYNDROME

Dysregulation of the HPA axis has been hypothesized to play a role in AD for several decades (Ahmad et al., 2019; Canet et al., 2018; Raskind et al., 1982; Spar & Gerner, 1982). The HPA axis is the biological system that controls the neuroendocrine response to perceived stressors. Broadly, in response to a stressor, the hypothalamus triggers the release of several hormones which subsequently trigger the secretion of ACTH into the bloodstream from the pituitary. This elevation stimulates the synthesis and secretion of cortisol, among other glucocorticoids (Herman et al., 2011). A portion of this secreted cortisol is then able to pass through the blood-brain barrier (Dunn et al., 1981; Pardridge & Mietus, 1979) where it can bind to mineralocorticoid (Type I) or glucocorticoid (Type II) receptors (Veldhuis et al., 1982). Binding in the hypothalamus directly inhibits further secretion while cortisol binding in several brain regions with high densities of mineralocorticoid and glucocorticoid receptors, most notably the medial prefrontal cortex and hippocampus (Klok et al., 2011; Seckl et al., 1991; Tohgi et al., 1995; Watzka et al., 2000; Webster et al., 2002) indirectly downregulates HPA axis activity (Herman et al., 2012; Kinlein & Karatsoreos, 2020).

Early in-vitro and rodent studies demonstrated that cortisol elevations, although not directly neurotoxic as originally hypothesized (Aus Der Mühlen & Ockenfels, 1968), decrease dendritic complexity (Magariños & McEwen, 1995; Magariños et al., 1998; Watanabe et al., 1992) and even neuron number (Sapolsky et al., 1985) and increase the vulnerability of hippocampal neurons to other insults such as excitotoxicity (Behl et al., 1997; Sapolsky, 1985). These findings, coupled with lesion studies demonstrating that
hippocampal damage can produce basal cortisol elevations (Bouillé & Baylé, 1973; Mandell et al., 1963; Rubin et al., 1966) led researchers to postulate that cortisol elevations may impair the very structures which downregulate cortisol. The hypothesized resultant feedback cycle was formalized by the GCH (Sapolsky et al., 1986). Although the GCH is not specific to AD, the application of this hypothesis and the role of the hippocampus in the regulation of the HPA axis has long been of particular interest in the context of AD (e.g., Jacobson & Sapolsky, 1991) due to the accelerated rates of hippocampal atrophy observed in AD (Jack et al., 1998; Zhao, W. et al., 2019) and accompanying impairment in hippocampal-dependent episodic memory (Lopez et al., 2011; Scoville & Milner, 1957).

Motivated in part by these models, numerous studies have measured cortisol in AD patients and controls (e.g., Gómez-Gallego & Gómez-García, 2018; Martignoni et al., 1990a; Pomara et al., 1984; Popp et al., 2009; Wang, L. Y. et al., 2018). While it is often tacitly assumed that AD patients have higher circulating cortisol levels than age-matched controls, a cursory examination of the literature reveals many studies showing no difference (e.g., Cascalheira et al., 2009; Parnetti et al., 1995; Porter et al., 2002), and even studies reporting numerically lower cortisol levels in AD than controls (e.g., Airaghi et al., 1991; de Bruin et al., 2002).

As is common with the examination of cortisol, the studies of cortisol levels in AD patients have been methodologically variable in the collection and measurement of cortisol. The literature includes comparisons of single or averaged measurements obtained from blood samples, area-under-the curve measurements derived from salivary samples taken over the course of several hours or whole days, and integrated measures of
cortisol from 24-hour urine samples. Additionally, participant characteristics such as the age and sex of the AD and control samples, the method of AD classification, and the severity of the clinical syndrome varied across-study. The number of studies and the methodological heterogeneity between studies examining cortisol levels in AD shows that a systematic quantitative review and meta-analysis of this literature was long overdue.

Therefore, this meta-analysis systematically evaluated the hypothesis that dysregulation of the HPA-axis is characteristic of AD. This was accomplished by comparing the basal cortisol levels of demented participants with those of cognitively normal older adults. Then, the potential moderating role of sample characteristics such as age, sex composition and dementia severity (as measured by MMSE (Folstein et al., 1975) performance) and methodological approaches such as bodily fluid used for cortisol collection or the specificity of AD diagnosis was examined.

Based on the GCH and evidence of hippocampal atrophy in AD patients, it was anticipated that basal cortisol levels would be higher in participants with AD compared to participants with normal cognition. The self-amplifying nature of the GCH also suggests that participants with more severe dementia would have higher cortisol levels than less severely impaired patients. Therefore, it was anticipated that effects determined from more severely impaired AD samples would be larger than those calculated from less severely impaired samples.

Reports of age-related cortisol elevations even in the absence of dementia (Deuschle et al., 1997; Nicolson et al., 1997; Purnell et al., 2004; Rueggeberg et al., 2012) led to the prediction that participant age would moderate this relationship. Specifically, it was anticipated that older samples would show greater differences in basal
cortisol levels than younger samples. Finally, larger effects were expected from samples including a larger proportion of female AD patients due to evidence of cortisol elevations in female AD compared to male AD patients (Leblhuber et al., 1993).
Method

Literature Search

A literature search was conducted using the PsychINFO and MEDLINE databases with the Boolean phrase (cortisol OR glucocorticoids) AND (Alzheimer’s OR dementia OR mild cognitive impairment OR MCI). The search was limited to articles published in peer-reviewed journals between 1/1/1984 and 12/31/2019, (a) written in English, (b) describing empirical studies (c) performed in living humans (d) including an adult population. This search resulted in 605 unique articles. See
Table 1 for clarification of the terminology used during article selection and the remaining analyses.
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assay</td>
<td>A single value obtained reflecting the cortisol content measured from a collection using a technique such as RIA or HPLC-MS. A single collection may be measured multiple times, resulting in multiple assays per collection. If this is done, it is typically done such that two assays are obtained per sample (i.e., “assays measured in duplicate”).</td>
<td>The result of a RIA test of a single 13:00h plasma collection</td>
</tr>
<tr>
<td>Collection</td>
<td>A quantity of fluid (e.g., a single blood draw, urine void, or saliva swab) obtained from a single participant at a single given time-point. Even if collections were obtained at nearly identical times (e.g., consecutive blood draws at 09:00 and 09:01), each separate fluid sample (e.g., unmixed vial) is considered a single collection</td>
<td>A swab of saliva obtained at 08:00 5 mL of blood obtained at 09:00</td>
</tr>
<tr>
<td>Occasion</td>
<td>A period of time over which one or more collections were obtained from a participant. Occasions are almost always synonymous with days.</td>
<td>A single day on which an 08:00 plasma collection was obtained</td>
</tr>
<tr>
<td>Index/indices</td>
<td>The outcome measure reported determined from one or more collections, obtained at one or more occasions.</td>
<td>The mean of the two assays from a saliva collection obtained at 09:00 and a saliva collection obtained at 12:00</td>
</tr>
<tr>
<td>Sample</td>
<td>A selection of participants grouped together on the basis of like traits. For this analysis, relevant traits are presence or absence of dementia and distribution of participant sex.</td>
<td>30 participants with AD 17 female participants with normal cognition</td>
</tr>
<tr>
<td>Sample pair</td>
<td>A set of two samples (one AD and one Control) used to calculate an effect size</td>
<td>A sample of 27 participants with AD and an age-matched sample of 26 control participants, each with an 09:00 cortisol collection</td>
</tr>
</tbody>
</table>
A sample of 10 male participants with AD and a sample of 12 male control participants, each with blood collections from 14:00 and 16:00.

<table>
<thead>
<tr>
<th>Effect</th>
<th>A single value reflecting the SMD (and corresponding SDs) obtained from two independent samples using identical collection and measurement methods.</th>
<th>The SMD compared between a sample of 14 AD participants and 13 cognitively normal participants.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>Refers to a protocol used for the designation of cognitive status sample membership, collection and measurement of cortisol. A single study may report multiple effects and a single article may include multiple studies.</td>
<td>The collection and subsequent measurement via RIA of 5 mL of blood at 09:00 from 43 participants who were tested using the MMSE.</td>
</tr>
<tr>
<td>Article</td>
<td>A single, peer-reviewed, published document.</td>
<td>(Ferrier et al., 1988)</td>
</tr>
</tbody>
</table>

*Note.* AUCg: Elements are arranged in order from the smallest to the largest units.
Each of the article abstracts was reviewed and tagged for full text review if it confirmed the literature search criteria, showed evidence of reporting the results of at least one study that included at least one demented population (of unspecified or AD specific etiology), and indicated that cortisol or other hormone was assayed from a biological collection.

The approach to the examination of full-text articles was intentionally very liberal, resulting in 209 full-text articles for further review. To be considered methodologically eligible for inclusion, at least one of the studies included in the article had to explicitly indicate that basal (i.e., unstimulated) cortisol was measured from a biological collection obtained and analyzed using identical protocols for a cognitively normal and a demented population.

All methodologically eligible articles were screened and articles reporting only duplicate samples were removed. To be included in the present analysis, the article had to report, at a minimum, (a) a conventional measure of central tendency (i.e., the mean or median) and corresponding spread (i.e., one of: the SD, SEM, a CI of the mean with specified confidence level, the range, IQR, or specification of the first and third quartiles) for an index of basal cortisol separately from both an eligible demented population and a control population, and (b) an indication of sample sizes for both the demented and control groups from the methodologically relevant study. All articles included in this analysis reported results from exactly one study, therefore the term study will be adopted with reference to information from included articles.

Studies were designated as incomplete if the measure of central tendency or spread was not reported for the cortisol indices using one of the conventional measures of
central tendency or spread but were referenced in the text (e.g., values appeared in a figure such that exact index values were not discernible or only the significance designation such as $p > 0.05$ was reported).
Figure 3. Process for sorting articles for inclusion in the present meta-analysis. For the Abstract Excluded and Full Text Excluded boxes numbers represent the number of articles excluded because they did not meet a given criteria. For example, 24 articles were excluded prior to reading the full text because the study or studies reported were not conducted using living persons. For all other boxes, numbers represent the number of articles that fit the relevant criteria (e.g., based on the reading of the abstracts 209 articles met the criteria to have the full text read).
Article Recording and Coding

Cognitive Status Grouping

The intention of this meta-analysis was to examine differences in basal cortisol levels between individuals with AD and healthy older adults. However, as briefly discussed in the Defining AD section, a distinction is made between AD and the clinical syndrome commonly associated with the disease. As the former requires biomarker analysis and is only indirectly assessed in vivo by the commonly used clinical diagnostic criteria (e.g., NINCDS-ADRDA, DSM, ICD) the analysis here is more accurately a comparison of the basal cortisol levels of demented and cognitively normal older adults.

For the purposes of this analysis, the cognitive designation of demented or normal made by the authors were used. If the etiology of dementia was specified as AD or was non-specific, the demented sample was labelled as AD here. No delineations of dementia severity (e.g., mild, moderate, severe) were used as these were only variably and inconsistently reported. However, independent samples specifically described as MCI were not included in the present analysis (i.e., they were not considered AD or NCC). Similarly, the author designation of normal cognition was accepted here without consideration of cognitive test performance. While the utility of subjective cognitive impairment is debated (see Burmester et al., 2016 for discussion), these participants were not considered NCC. Finally, some studies contained both young and older NCC groups. Younger NCC samples were excluded if the central tendency and spread measures were available for the older NCC sample.
Cortisol Indices and Covariates

For all included studies, all reported cortisol indices were recorded as well as details regarding the timing of fluid collection and whether duplicate assays were conducted. Although some cortisol research has focused on the interpretation of ratio or rate indices, such comparisons of assays from evening and morning collections (e.g., Barca et al., 2019; Ferrari et al., 2000) or the CAR (Clow et al., 2004) (e.g., Dijckmans et al., 2017), these indices were not considered.

In addition to the sample sizes and cortisol indices several covariates were recorded and coded as follows:

(a) Fluid: The biological fluid (i.e., blood, saliva, or urine) collected for cortisol assay.
(b) Age: Measures of central tendency or spread of the age of the participants.
(c) Sex: The number of male and female participants in each sample.
(d) MMSE Score: Measures of central tendency or spread of the MMSE scores of each sample.

For all covariates, missing information was imputed whenever possible. Specifically, if samples were reported as matched, exactly matched values were imputed where missing. Similarly, if only a subset of participants for which demographics were reported had cortisol indices, the cortisol subset was assumed to follow the demographic distribution of the larger sample. Finally, if an article referenced a previous source or article containing more information about the sample(s) or protocols, these sources were examined and any relevant covariate information that could be gleaned was added.
**Data Preparation**

Most studies included the mean (\(x\)) and SD (\(s\)) for reported cortisol indices and relevant covariates. The relevant effect size was calculated from these values directly. If other metrics of central tendency and/or spread were reported for cortisol indices or covariates these metrics were converted to means and SDs as accurately as possible. If the SEM or a CI was reported these values were mathematically transformed to SD units (Higgins et al., 2019). If the median and range or IQR were reported, the mean was estimated using the strategy presented by Luo et al. (2018) and corresponding SD estimated as described by Wan et al. (2014).

**Effect Size Estimation**

SMD measures were used to compare cortisol levels between older adults with AD and NCC adults. The SMD is the difference in group means when all within-group SDs are transformed to one (Borenstein, 2009). All differences (SMD and difference moderators) were calculated as AD minus control. For each sample pair of means, SD, and sample sizes, the SMD (\(d\)) was calculated with adjustment for positive bias using Hedges \(g\) (Hedges, 1981), such that the \(d\) for effect \(i\) is:

\[
d_i = \left[1 - \frac{3}{4(n_{Ai} + n_{Ci} - 9)}\right] \frac{\bar{x}_{Ai} - \bar{x}_{Ci}}{S_i}
\]

(1)

Where the \(A\) and \(C\) subscripts reflect the parameters of the AD and NCC samples respectively, such that positive values of \(d\) reflect higher cortisol levels in the AD compared to the NCC sample. The pooled SD (\(S_i\)) is calculated as:

\[
S_i = \sqrt{\frac{(n_{Ai} - 1)s_{Ai}^2 + (n_{Ci} - 1)s_{Ci}^2}{n_{Ai} + n_{Ci} - 2}}
\]

(2)
Where $S^2_{AI}$ and $S^2_{CI}$ are the AD and NCC sample variances. Finally, the within-effect variance of $d_i$ (correcting for small sample bias) (Hedges & Olkin, 1985; Valentine et al., 2009, p. 264) was estimated as:

$$\hat{\sigma}_i^2 = \frac{n_{AI} + n_{CI}}{n_{AI}n_{CI}} + \frac{d_i^2}{2(n_{AI} + n_{CI})}$$

### Effect Size Independence

If a single study reported multiple indices, a single effect size was selected or calculated per sample pair with the intention of maximizing the amount of independent information included in the analysis. If a study reported a single summary index which included the average of other reported indices the summary measure which included the largest number of samples (and corresponding SDs) was used. For example, Bemelmans et al. (2007) reported sample means and SDs for plasma cortisol levels collected at 09:00, 12:00 and 16:00, as well as the AUC calculated using all three indices. For this analysis, only the AUC values were used. If multiple non-independent summary measures were reported, the measure with the most samples was included. For example, Dijckmans et al. (2017) reported sufficient information for an index that consisted of the mean of assays from collections at waking, 30-minutes after waking, 12:00, 17:00, and 21:00, as well as the mean of assays from collections taken at 12:00, 17:00, and 21:00. Only the former index was included.

If multiple independent indices were reported, all independent cortisol indices were averaged\(^2\). This averaged value (and corresponding SD) was used to calculate the

---

\(^2\) With the exception of the sample pair from Popp et al. (2015), see Fluid section of the Discussion.
effect size (i.e., the SMD) for the sample pair. Finally, if the cortisol indices were reported separately for males and females in both the AD and NCC samples, relevant information for all four samples was recorded and two effects were calculated, one from each of the two sample pairs. *Table 2* shows the cortisol values for all included effects and covariate information as grouped by the article reporting the relevant effect(s).
Table 2: Included Effects and Covariates by Article

<table>
<thead>
<tr>
<th>Authors</th>
<th>#</th>
<th>N (Female)</th>
<th>Age (AD)</th>
<th>Age (NC)</th>
<th>Cortisol (AD)</th>
<th>Cortisol (NC)</th>
<th>MMSE (AD)</th>
<th>MMSE (NC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airaghi et al. (1991)</td>
<td>1</td>
<td>14 (10)</td>
<td>13 (9)</td>
<td>72.6 (5.9)</td>
<td>14.3 (1.6)</td>
<td>72.7 (6)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Armanini et al. (2003)</td>
<td>2</td>
<td>23 (18)</td>
<td>23 (18)</td>
<td>66.7 (8.8)</td>
<td>0.8 (0.2)</td>
<td>66.8 (10.4)</td>
<td>19.9 (2.3)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Arsenault-Lapierre et al. (2010)</td>
<td>3</td>
<td>12 (5)</td>
<td>24 (13)</td>
<td>80.2 (4.5)</td>
<td>0.3 (0.2)</td>
<td>75.8 (8)</td>
<td>20.1 (4.2)</td>
<td>28.8 (1.2)</td>
</tr>
<tr>
<td>Barca et al. (2019)</td>
<td>4</td>
<td>319 (184)</td>
<td>124 (61)</td>
<td>79.1 (9.2)</td>
<td>14.2 (41.8)</td>
<td>68.8 (11.5)</td>
<td>18.2 (6.3)</td>
<td>28.5 (2)</td>
</tr>
<tr>
<td>Bemelmans et al. (2007)</td>
<td>5</td>
<td>21 (18)</td>
<td>21 (13)</td>
<td>85.2 (5.1)</td>
<td>18.3 (6.2)</td>
<td>85.7 (6.4)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Cascalheira et al. (2009)</td>
<td>6</td>
<td>9 (9)</td>
<td>18 (18)</td>
<td>74.4 (16.3)</td>
<td>15.6 (8.4)</td>
<td>69.8 (17.3)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Chang et al. (2018)</td>
<td>8</td>
<td>21 (14)</td>
<td>20 (11)</td>
<td>78.5 (5.7)</td>
<td>10.2 (2.9)</td>
<td>77.5 (4.4)</td>
<td>16.1 (5.4)</td>
<td>27.5 (1.9)</td>
</tr>
<tr>
<td>Csermanskly et al. (2006)</td>
<td>9</td>
<td>10 (6)</td>
<td>21 (14)</td>
<td>76 (4.6)</td>
<td>18 (4.3)</td>
<td>77.6 (9.7)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Cunningham et al. (2001)</td>
<td>10</td>
<td>52 (52)</td>
<td>60 (60)</td>
<td>77.1 (6)</td>
<td>96.6 (4.8)</td>
<td>69.8 (6.3)</td>
<td>335 (77.7)</td>
<td>28.7 (1.7)</td>
</tr>
<tr>
<td>Curto et al. (2017)</td>
<td>11</td>
<td>18 (11)</td>
<td>22 (12)</td>
<td>80.1 (7.9)</td>
<td>351 (246)</td>
<td>74.1 (14.5)</td>
<td>13 (9.5)</td>
<td>26.3 (3.3)</td>
</tr>
<tr>
<td>Davis, K. L. et al. (1986)</td>
<td>12</td>
<td>23 (8)</td>
<td>9 (3)</td>
<td>66.6 (7.9)</td>
<td>9.6 (4.8)</td>
<td>59.1 (7.2)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>de Bruin et al. (2002)</td>
<td>13</td>
<td>16 (16)</td>
<td>17 (17)</td>
<td>79 (4.3)</td>
<td>13.8 (4.1)</td>
<td>76 (4.7)</td>
<td>16.8 (3.3)</td>
<td>25.4 (2.4)</td>
</tr>
<tr>
<td>de la Rubia Ortí et al. (2017)</td>
<td>15</td>
<td>20 (-)</td>
<td>20 (-)</td>
<td>73.6 (-)</td>
<td>18.6 (4.1)</td>
<td>73.6 (-)</td>
<td>16.8 (3)</td>
<td>25.4 (3.2)</td>
</tr>
<tr>
<td>Ferrier et al. (1988)</td>
<td>16</td>
<td>15 (13)</td>
<td>16 (10)</td>
<td>80 (8)</td>
<td>4.5 (2)</td>
<td>75 (6)</td>
<td>356 (147)</td>
<td>304 (98)</td>
</tr>
<tr>
<td>Franceschi et al. (1991)</td>
<td>17</td>
<td>14 (10)</td>
<td>13 (10)</td>
<td>72.6 (5.9)</td>
<td>15.5 (5)</td>
<td>72.7 (6)</td>
<td>18 (6.5)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Giubilei et al. (2001)</td>
<td>18</td>
<td>18 (7)</td>
<td>18 (7)</td>
<td>66.9 (6.2)</td>
<td>10.6 (8.1)</td>
<td>61.1 (5)</td>
<td>18 (6.5)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Gómez-Gallego and Gómez-Garcia</td>
<td>19</td>
<td>46 (32)</td>
<td>52 (36)</td>
<td>78.3 (4.8)</td>
<td>13.5 (7.5)</td>
<td>76.7 (7.5)</td>
<td>18.8 (4.9)</td>
<td>28.6 (1.5)</td>
</tr>
<tr>
<td>(2018)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gomez-Gallego and Gomez-Garcia</td>
<td>20</td>
<td>80 (56)</td>
<td>104 (72)</td>
<td>79.4 (8.7)</td>
<td>6.6 (0.4)</td>
<td>74 (7.5)</td>
<td>18.9 (4.9)</td>
<td>28.5 (1.5)</td>
</tr>
<tr>
<td>(2019)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hartmann et al. (1997)</td>
<td>21</td>
<td>12 (8)</td>
<td>10 (3)</td>
<td>63 (8.4)</td>
<td>256 (92.2)</td>
<td>68 (8.4)</td>
<td>19 (6.3)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Higuchi et al. (2010)</td>
<td>22</td>
<td>16 (12)</td>
<td>16 (12)</td>
<td>73.4 (5.2)</td>
<td>23.6 (5)</td>
<td>74.1 (3.8)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Authors</td>
<td>#</td>
<td>N (Female)</td>
<td>Age</td>
<td>Cortisol</td>
<td>MMSE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>----</td>
<td>------------</td>
<td>-----</td>
<td>----------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total (AD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AD</td>
<td>NC</td>
<td>AD</td>
<td>NC</td>
<td>AD</td>
<td>NC</td>
</tr>
<tr>
<td>James et al. (2019)</td>
<td>23</td>
<td>65 (43)</td>
<td>76.5 (7.9)</td>
<td>71 (8.7)</td>
<td>0.8 (0.3)</td>
<td>0.8 (0.3)</td>
<td>21 (2.4)</td>
<td>28.1 (2.4)</td>
</tr>
<tr>
<td>Johar et al. (2015)</td>
<td>24</td>
<td>33 (10) 599 (318)</td>
<td>77.8 (6.8)</td>
<td>74.4 (6.1)</td>
<td>0.8 (1.1)</td>
<td>0.9 (1.2)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Laske et al. (2009)</td>
<td>25</td>
<td>26 (16) 20 (7)</td>
<td>70.9 (8.1)</td>
<td>70 (11.8)</td>
<td>551 (146)</td>
<td>435 (83.9)</td>
<td>23.5 (1.6)</td>
<td>28.5 (1.5)</td>
</tr>
<tr>
<td>Laske et al. (2011)</td>
<td>26</td>
<td>85 (52) 70 (39)</td>
<td>72.4 (23.5)</td>
<td>70.8 (7.9)</td>
<td>486 (139)</td>
<td>365 (142)</td>
<td>21.3 (3.7)</td>
<td>29.2 (0.7)</td>
</tr>
<tr>
<td>Leake et al. (1990)</td>
<td>27</td>
<td>11 (9) 9 (6)</td>
<td>77 (6.3)</td>
<td>73 (4.8)</td>
<td>426 (224)</td>
<td>278 (78)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Leblhuber et al. (1993)</td>
<td>28</td>
<td>11 (11) 10 (10)</td>
<td>78.8 (5)</td>
<td>75.3 (7)</td>
<td>24.4 (6)</td>
<td>17.1 (4.6)</td>
<td>5.2 (6.5)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Martignoni et al. (1990a)</td>
<td>29</td>
<td>13 (0) 10 (0)</td>
<td>76.3 (7.1)</td>
<td>75.3 (8.6)</td>
<td>18.6 (4.1)</td>
<td>16.5 (4.5)</td>
<td>5.2 (6.5)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Martignoni et al. (1990b)</td>
<td>30</td>
<td>10 (2) 11 (4)</td>
<td>63.8 (9.7)</td>
<td>64.5 (6.4)</td>
<td>161 (53.1)</td>
<td>99.1 (30.5)</td>
<td>16 (5.2)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Masera et al. (2002)</td>
<td>31</td>
<td>8 (1) 8 (1)</td>
<td>65.8 (8.2)</td>
<td>65.2 (0.4)</td>
<td>95.8 (58.3)</td>
<td>84.7 (33.1)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Masugl et al. (1989)</td>
<td>32</td>
<td>16 (6) 17 (6)</td>
<td>66.9 (7.6)</td>
<td>66.7 (7.8)</td>
<td>13.3 (2.8)</td>
<td>10.2 (4.9)</td>
<td>19.6 (2.4)</td>
<td>29.6 (0.8)</td>
</tr>
<tr>
<td>Miller et al. (1994)</td>
<td>33</td>
<td>10 (0) 10 (0)</td>
<td>77.5 (5.5)</td>
<td>76.6 (6.8)</td>
<td>17.3 (4.3)</td>
<td>12 (3.1)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Molchan et al. (1990)</td>
<td>34</td>
<td>23 (7) 13 (2)</td>
<td>67.5 (7.2)</td>
<td>65.9 (7.7)</td>
<td>11.5 (1.9)</td>
<td>7.4 (2.2)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Muralado et al. (1993)</td>
<td>35</td>
<td>34 (19) 14 (3)</td>
<td>63.7 (9.5)</td>
<td>62.2 (19.2)</td>
<td>15.1 (5.3)</td>
<td>10.8 (4.6)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Muralado et al. (2000)</td>
<td>36</td>
<td>7 (6) 8 (5)</td>
<td>64.8 (10.6)</td>
<td>68.7 (11)</td>
<td>204 (16.9)</td>
<td>131 (6.5)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Närman et al. (1995b)</td>
<td>37</td>
<td>14 (11) 12 (9)</td>
<td>72.1 (6.7)</td>
<td>73 (5.1)</td>
<td>392 (123)</td>
<td>222 (160)</td>
<td>17.6 (6.7)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Närman et al. (1995a)</td>
<td>38</td>
<td>35 (26) 20 (10)</td>
<td>76.8 (12.2)</td>
<td>73.8 (14.5)</td>
<td>432 (285)</td>
<td>437 (209)</td>
<td>18.3 (9.3)</td>
<td>29.4 (1.3)</td>
</tr>
<tr>
<td>Násman et al. (1995a)</td>
<td>39</td>
<td>10 (10) 9 (9)</td>
<td>74.6 (6.5)</td>
<td>74.2 (7.6)</td>
<td>373 (104)</td>
<td>342 (51)</td>
<td>18.5 (5.5)</td>
<td>29.4 (0.7)</td>
</tr>
<tr>
<td>Násman et al. (1995b)</td>
<td>40</td>
<td>8 (0) 10 (0)</td>
<td>74.6 (6.5)</td>
<td>74.2 (7.6)</td>
<td>378 (86)</td>
<td>371 (101)</td>
<td>18.5 (5.9)</td>
<td>29.4 (0.6)</td>
</tr>
<tr>
<td>Násman et al. (1996)</td>
<td>41</td>
<td>23 (15) 19 (9)</td>
<td>74.2 (7.4)</td>
<td>74.2 (7.7)</td>
<td>253 (110)</td>
<td>218 (94)</td>
<td>18.3 (4.4)</td>
<td>29.4 (0.5)</td>
</tr>
<tr>
<td>Parnetti et al. (1990)</td>
<td>42</td>
<td>21 (6) 41 (19)</td>
<td>72.5 (9.8)</td>
<td>71.7 (10.2)</td>
<td>7.2 (4)</td>
<td>7.6 (7.2)</td>
<td>17.9 (6.2)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Parnetti et al. (1995)</td>
<td>43</td>
<td>32 (14) 44 (22)</td>
<td>73.9 (1.6)</td>
<td>72.2 (1.5)</td>
<td>7 (4.3)</td>
<td>7.1 (3.7)</td>
<td>13.1 (1.9)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Peabody et al. (1986)</td>
<td>44</td>
<td>9 (0) 9 (0)</td>
<td>63.4 (6.4)</td>
<td>63.8 (7)</td>
<td>13.5 (2.8)</td>
<td>12.6 (3.8)</td>
<td>13.5 (8.2)</td>
<td>28.4 (1.4)</td>
</tr>
<tr>
<td>Pomara et al. (1984)</td>
<td>45</td>
<td>60 (-) 14 (-)</td>
<td>70.7 (8.6)</td>
<td>69.6 (9)</td>
<td>13.6 (3.9)</td>
<td>13.3 (4.9)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Pomara et al. (1988)</td>
<td>46</td>
<td>10 (7) 10 (6)</td>
<td>66.4 (6.4)</td>
<td>69.6 (8.5)</td>
<td>13.5 (3.9)</td>
<td>16.9 (7.3)</td>
<td>- (-)</td>
<td>- (-)</td>
</tr>
<tr>
<td>Popp et al. (2015)</td>
<td>47</td>
<td>105 (63) 37 (16)</td>
<td>73 (7.4)</td>
<td>64.3 (8.1)</td>
<td>15.2 (6.2)</td>
<td>16.3 (7.6)</td>
<td>22.7 (3.1)</td>
<td>28.7 (1.1)</td>
</tr>
<tr>
<td>Authors</td>
<td>N</td>
<td>Total (Female)</td>
<td>Mean (SD)</td>
<td>Cortisol</td>
<td>MMSE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----</td>
<td>----------------</td>
<td>-----------</td>
<td>----------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AD</td>
<td>NC</td>
<td>AD</td>
<td>NC</td>
<td>AD</td>
<td>NC</td>
<td>AD</td>
</tr>
<tr>
<td>Rasmuson et al. (1998)</td>
<td>48</td>
<td>8</td>
<td>(8)</td>
<td>7</td>
<td>(7)</td>
<td>78</td>
<td>(8.4)</td>
<td>76.7</td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>5</td>
<td>(0)</td>
<td>5</td>
<td>(0)</td>
<td>78</td>
<td>(8.4)</td>
<td>76.7</td>
</tr>
<tr>
<td>Rasmuson et al. (2001)</td>
<td>50</td>
<td>10</td>
<td>(10)</td>
<td>7</td>
<td>(7)</td>
<td>80</td>
<td>(5.8)</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>12</td>
<td>(0)</td>
<td>10</td>
<td>(0)</td>
<td>76.4</td>
<td>(7.8)</td>
<td>75.4</td>
</tr>
<tr>
<td>Rasmuson et al. (2002)</td>
<td>51</td>
<td>21</td>
<td>(21)</td>
<td>12</td>
<td>(12)</td>
<td>76.4</td>
<td>(7.8)</td>
<td>75.4</td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>12</td>
<td>(0)</td>
<td>10</td>
<td>(0)</td>
<td>76.4</td>
<td>(7.8)</td>
<td>75.4</td>
</tr>
<tr>
<td>Rolandi et al. (1992)</td>
<td>53</td>
<td>6</td>
<td>(0)</td>
<td>6</td>
<td>(0)</td>
<td>70</td>
<td>(1.6)</td>
<td>82.5</td>
</tr>
<tr>
<td>Spada et al. (2001)</td>
<td>54</td>
<td>15</td>
<td>(7)</td>
<td>23</td>
<td>(8)</td>
<td>70</td>
<td>(8)</td>
<td>68</td>
</tr>
<tr>
<td>Swanwick et al. (1998)</td>
<td>55</td>
<td>18</td>
<td>(11)</td>
<td>17</td>
<td>(9)</td>
<td>74.4</td>
<td>(4.7)</td>
<td>70.5</td>
</tr>
<tr>
<td>Tollefson et al. (1989)</td>
<td>56</td>
<td>25</td>
<td>(20)</td>
<td>76</td>
<td>(62)</td>
<td>72.1</td>
<td>(9.9)</td>
<td>71.9</td>
</tr>
<tr>
<td>Tsuboyama et al. (1992)</td>
<td>57</td>
<td>28</td>
<td>(10)</td>
<td>19</td>
<td>(4)</td>
<td>67.2</td>
<td>(7.4)</td>
<td>65.9</td>
</tr>
<tr>
<td>Umegaki et al. (2000)</td>
<td>58</td>
<td>66</td>
<td>(66)</td>
<td>21</td>
<td>(21)</td>
<td>82.5</td>
<td>(7.8)</td>
<td>83</td>
</tr>
<tr>
<td>Vankova et al. (2016)</td>
<td>59</td>
<td>16</td>
<td>(16)</td>
<td>22</td>
<td>(22)</td>
<td>74.8</td>
<td>(10.5)</td>
<td>66.9</td>
</tr>
<tr>
<td>Wirth et al. (2019)</td>
<td>60</td>
<td>112</td>
<td>(47)</td>
<td>58</td>
<td>(28)</td>
<td>74.8</td>
<td>(8)</td>
<td>75.1</td>
</tr>
<tr>
<td>Zvěřová et al. (2013)</td>
<td>61</td>
<td>45</td>
<td>(28)</td>
<td>37</td>
<td>(29)</td>
<td>75.1</td>
<td>(8.4)</td>
<td>63.2</td>
</tr>
<tr>
<td>BLSA</td>
<td>62</td>
<td>47</td>
<td>(47)</td>
<td>47</td>
<td>(47)</td>
<td>84.3</td>
<td>(7.2)</td>
<td>82.8</td>
</tr>
<tr>
<td></td>
<td>63</td>
<td>54</td>
<td>(0)</td>
<td>54</td>
<td>(0)</td>
<td>84.5</td>
<td>(4.9)</td>
<td>84.2</td>
</tr>
</tbody>
</table>

*Note.* # refers to the effect number. This is a unique identifier assigned to each independent effect included in at least one of the analyses conducted as a part of this meta-analysis. Articles included in Analyses One through Five are marked with an * in the References.

1Mean estimated using reported range 2Mean estimated using reported median and range 3Mean estimated using reported median and Q1, Q3 and/or IQR 4SD estimated using reported SEM 5SD estimated using reported CI 6SD estimated using reported range 7SD estimated using reported Q1, Q3 and/or IQR.
Addition of the BLSA Data

Because Ennis et al. (2017) only reported cortisol levels measured prior to dementia onset for the AD sample, the BLSA cortisol measures could not be included as published in that article. However, the authors generously allowed full access to their database, allowing for the comparison of cortisol levels between AD and NCC in the BLSA included here.

Due to the longitudinal nature of the study, most participants have multiple cortisol measures. Of 1,865 participants contributing cortisol collections, 1,511 participants remained dementia free through their last BLSA visit and contributed a total of 4,711 cortisol collections. One-hundred-and-one participants diagnosed with AD contributed a total of 173 cortisol collections after AD onset. To maximize the similarity between the included BLSA data and the data incorporated from the other studies in this meta-analysis, only the cortisol level from one collection was incorporated for each participant. For the AD participants, the cortisol level from the last cortisol collection was included.

After selecting a single cortisol collection for each AD participant, a 1:1 sample matching was done using all 4,711 cortisol collections obtained from NCC participants via the MatchIt package (Ho et al., 2007; Ho et al., 2011) with nearest neighbor matching on the basis of age at collection, sex, and APOE genotype. The default match procedure was modified so that once a single cortisol collection occasion was matched, all collection contributed by that participant were removed from the pool of potential matches. This insured that no participant was represented in the NCC sample twice and resulted in a sample of 202 participants used to calculate Effects 62 and 63 (see Table 2).
Analytic Approach

All analyses were conducted using the metafor package (Viechtbauer, 2010) in R version 4.0.1 (R Core Team, 2020). First, a single three-level (Van Den Noortgate & Onghena, 2003) random-effects model (Hedges & Vevea, 1998) was fit to all effects meeting inclusion criteria, using study as the clustering factor for the three-level approach. Next, an analogous random-effects model excluding outlying and influential effects and effects missing covariate information was applied to eligible effects. Finally, a series of mixed-effects models to examine whether between-study heterogeneity could be explained by various moderators. The restricted maximum likelihood estimator (Viechtbauer, 2005, 2010), a largely unbiased estimator of between-effect variance (Langan et al., 2019), was used for all models.

Data Examination

The balance of AD and NCC participants in each sample pair was examined as a function of total sample pair size. As shown in Figure 4, the average sample pair size was 65 participants and sample pairs were approximately balanced (i.e., included approximately equal AD and NCC sample sizes). However, it is noteworthy that the two largest sample sizes (both more than six times larger than the average paired sample size) are largely imbalanced.

As shown in Figure 5, in line with the random-effects model assumption, the effect sizes followed a normal distribution (with the exception of Effect 36). As anticipated, there was a log-linear relationship between effect size variance and sample pair size (see Figure 6), again with the exception of Effect 36. Finally, there was no evidence of an effect of year of publication on effect size (see Figure 7).
Figure 4. The ratio of AD and NCC participants used for effect size calculation. Each point reflects an effect size. The solid line indicates a sample size ratio of 1 (i.e., the AD and NCC sample sizes are equivalent), while points above the line reflect sample pairs with a greater portion of AD participants than NCC participants. The red point reflects the mean ratio and the mean sample pair size. Labels reflect the effect number corresponding to each sample pair.

Figure 5. The distribution of the primary study effect sizes. Points reflect a single SMD between the AD and NCC samples. Darker colors indicate a smaller sample size while lighter colors indicate a larger sample size, with the lightest color representing sample sizes greater than 200 participants.
Figure 6. The relationship between sample pair size \((N)\) and the variance of the effect size. The expected negative relationship between sample size and variance is present, such that effect sizes determined from larger sample sizes have less variance. However, the X-axis scale is logarithmic rather than linear, demonstrating that the rate of decline in variance slows as the sample pair size increases.

Figure 7. Scatterplot of the relationship between the SMD of cortisol levels for the AD and NCC samples as a function of year of publication.
**Meta-Analytic Analyses**

**Analysis One** examined the SMD between basal cortisol indices reported for AD and NCC samples. All possible effect sizes were included in Analysis One, regardless of covariate information, outlier, or influence status. The resulting Cook’s distance (Cook & Weisberg, 1982) and DFBETAS (Viechtbauer & Cheung, 2010) values were calculated for each effect to examine influence. **Analysis Two** replicated Analysis One, excluding several outlying and influential points identified using the DFBETAS and Cook’s distance values from Analysis One. **Analysis Three** examined the influence of the collection fluid (blood, saliva, or urine) on the SMD between AD and NCC samples.

**Analysis Four** examined the influence of age and sex, on the effect size. Two covariate values were used to examine the influence of age: the weighted average age of the sample pair and the age difference in the samples. The age difference was calculated as the average age of the AD sample minus the average age of the NCC sample from each sample pair. Therefore, positive values reflect an older AD than NCC sample. Similarly, the weighted average proportion of female participants in each sample and the difference in the proportion of female participants between the AD and NCC samples for each sample pair were included as covariates. Again, the difference was calculated as AD minus NCC. Finally, **Analysis Five** examined the influence of dementia severity (indexed by the AD sample mean score on the MMSE) on the SMD of cortisol between AD and NCCs.
Results

Analysis One: Basal Cortisol Differences between AD and NCC

Analysis One included all 63 effects shown in Table 2. Without adjustment for covariates or consideration of outlying or influential points, $d = 0.46$ ($p < .001$) demonstrating that the reported cortisol indices were an average of approximately half of a SD higher in the AD than the NCC samples. The results of Analysis One also reflect the anticipated significant amount of between-study heterogeneity (see Table 3).

The influence diagnostics (Cook’s Distance and DFBETAS, shown in Figure 8) demonstrated that there were several influential effects, namely Effects 36, 2, and 34. Although the Cook’s Distances for these effects are relatively small (well under the 0.45 rule of thumb (Viechtbauer & Cheung, 2010)) they are clearly distinct from the remaining Cook’s Distance values. Similarly, examining the DFBETAS values, these three effects stand out from the others. Therefore, these effects were not included in subsequent analyses.
Table 3: Meta-Analytic Summary Results

<table>
<thead>
<tr>
<th>k</th>
<th>Q_M</th>
<th>Q_E</th>
<th>$\tau^2$</th>
<th>b</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>63</td>
<td>37.769***</td>
<td>213.957***</td>
<td>0.211</td>
<td>Intercept 0.464*** ( 0.316 - 0.613 )</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>34.731***</td>
<td>160.373***</td>
<td>0.131</td>
<td>Intercept 0.388*** ( 0.259 - 0.517 )</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>0.162</td>
<td>158.826***</td>
<td>0.136</td>
<td>Intercept 0.447** ( 0.133 - 0.762 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.136</td>
<td>Fluid 0.044 ( -0.258 - 0.170 )</td>
</tr>
<tr>
<td>4</td>
<td>57</td>
<td>4.943</td>
<td>125.798***</td>
<td>0.106</td>
<td>Intercept 0.706 ( -1.188 - 2.600 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.004</td>
<td>Age (w) -0.004 ( -0.030 - 0.022 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.695</td>
<td>Female (d) 0.695 ( -0.321 - 1.712 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.015</td>
<td>Female (w) 0.015 ( -0.315 - 0.345 )</td>
</tr>
<tr>
<td>5</td>
<td>39</td>
<td>1.861</td>
<td>113.520***</td>
<td>0.124</td>
<td>Intercept 0.861* ( 0.160 - 1.562 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.124</td>
<td>MMSE -0.026 ( -0.064 - 0.011 )</td>
</tr>
</tbody>
</table>

**Note.** b: regression coefficient; covariates. (d): Represents the regression coefficient for the difference in the AD and control groups; k reflects the number of effects included; (w) represents the regression coefficient for the weighted average of the AD and control groups. Q_M tests whether all regression coefficients (b) are simultaneously zero; Q_E tests whether significant between-study heterogeneity ($\tau^2$). The 95% CI reflects the CI of the regression coefficients.

***p < 0.001 **p < 0.01 *p < 0.05
Figure 8. The influence diagnostics from Analysis One. Each point reflects an effect and the corresponding size represents the inverse variance of that effect. The x-axis represents the Cook’s distance between the estimated overall effect size with the $ith$ effect excluded. Larger values indicate a more influential study. The y-axis represents the DFBETAS value with larger absolute values reflecting more influence. Effects that were excluded from Analyses Two through Five are highlighted in red. The labels identify the effect number corresponding to each point.
Analysis Two: Elimination of Influential Effects

Analysis Two replicated Analysis One excluding the highly influential and outlying effects. This resulted in a random-effects analysis of 60 effects. Like Analysis One, the weighted SMD was significantly different than zero \( (d = 0.39, Q_M = 34.73, p < 0.001) \) and there was still significant heterogeneity between effects \( (\tau^2 = 0.13, Q_E = 160.37, p < 0.001) \). However, this reflects a large reduction in between-study variance (from \( \tau^2 = 0.21 \) in Analysis One), suggesting that the removal of the influential points resulted in a more homogenous set of effects. Further, the funnel plot (Sterne et al., 2011) shown in Figure 9 shows that effects are fairly normally distributed with only a handful of studies falling outside of the pseudo 95% confidence interval.
Figure 9. The funnel plot for the effects included in Analyses Two through Four. Each point represents an effect and the dotted black line reflects a pseudo 95% confidence interval.
Analysis Three: Effect of Fluid Type

Fluid type was not a significant moderator of effect size \((Q_M = 0.16, p = 0.68)\) and significant residual between-study heterogeneity remained in Analysis Three \((Q_E = 158.83, p < 0.001)\). Post-hoc examination of the average effect sizes for each of the three fluids separately (shown in Figure 10) revealed that the majority (47/60) of effects were calculated from blood collections and the results for this subset mirror the results from Analysis Two \((d = 0.42, Q_M = 32.58, p < 0.01)\). A smaller number of effects (8/60) were calculated using salivary collections, and the average of these effects was numerically smaller than the mean effect and was not significantly different than zero \((d = 0.18, Q_M = 1.80, p = 0.18)\). Finally, the mean of the five of 60 effects calculated from urine samples was not significantly different than zero \((d = 0.51, Q_M = 1.39, p = 0.24)\).
Figure 10. Forest plot of the effects included in Analyses Two and Analysis Three. Squares (sized by weight) reflect the SMD, effect numbers (see Table 2) are identified in the left-hand column. Bars identify the 95% CI (also shown in brackets on the right-most column). The x-axis reflects the SMD such that the vertical dotted line at 0 reflects no difference in cortisol levels between the AD and NCC samples.
Analysis Four: Effect of Sample Demographics

Analysis Four examined the influence of four demographic covariates of interest (the weighted mean and the difference for both sample ages and proportion of female participants in each sample). None of these covariates accounted for a significant amount of between-effect variance ($ps > 0.12$), nor was the model including all five covariates significant ($Q_M = 4.95, p = 0.29$). Further there was still a significant amount of residual heterogeneity between-effects after the consideration of the covariates ($Q_E = 125.80, p < 0.001$). See Table 3. The relationships between the effect sizes and the covariates used in Analysis Three are displayed in Figure 11.
Figure 11. The relationships between the weighted mean covariates of the sample pairs and the SMD of cortisol levels between the AD and NCC samples. Each point reflects an effect size and the length of each line reflects the raw difference between the mean of the AD sample and the mean of the NCC sample. Red points and lines represent those sample pairs for which the mean age or percentage of female participants was higher in the AD sample than the NCC sample. Black points and lines represent those sample pairs for which the NCC sample was older or a higher percentage of female participants than the AD sample. Gray points reflect sample pairs for which there was no difference in the mean age or female percentage values between sample pairs. Panel A shows the relationship between the SMD and the weighted mean age of the sample pairs. Panel B shows the relationship between the SMD and the weighted mean percentage of the sample pair that was female. For panel B, the gray dots at either end reflect effects that were calculated from entirely male or entirely female sample pairs, while the gray dots towards the center of the plot reflect estimation from matching.
Analysis Five: Effect of Dementia Severity

The last analysis examined the influence of the AD sample MMSE score on the SMD of cortisol levels between the AD and NCC samples. As shown in Table 3, the MMSE score did not explain a significant portion of the between-effect variation ($Q_M = 1.86, p = 0.17$). Figure 12 graphically depicts this relationship.
Figure 12. Scatterplot showing the relationship between the average MMSE score for the AD sample and the SMD of cortisol for the AD and NCC samples. The solid line represents the value (i.e., slope) of the regression coefficient for the MMSE score and the dotted lines show the 95% confidence interval for estimates at each possible value of MMSE score. The relationship is numerically negative \((b = -0.026)\) but insignificant \((p = 0.17)\).
Discussion

The combined results of Analyses One through Five show that older adults with AD have basal cortisol levels that are almost one half of a SD higher than those of older adults with normal cognitive status. The results of this formal meta-analysis are consistent with models and hypotheses reported in the literature of dysregulation of the HPA axis in AD (Ahmad et al., 2019; Canet et al., 2018; Jacobson & Sapolsky, 1991). At the most basic level, these theories propose that prolonged cortisol elevations damage the neurons of the hippocampus, one of the structures involved in the downregulation of further cortisol secretion (Sapolsky et al., 1986). This is consistent with evidence of advanced hippocampal atrophy and impairment on hippocampal-dependent functions observed in AD patients (Jack et al., 1998; Lopez et al., 2011; Zhao, W. et al., 2019) and the confirmation of basal cortisol elevations in AD reported here.

This meta-analysis did not attempt to determine the timing of the initiation of HPA-axis dysregulation in AD. However, the cyclical nature of damage to the hippocampus and subsequent further HPA-axis dysregulation coupled with the lack of clear restriction of hypercortisolism to a specific point in the disease progression suggest that targeting glucocorticoid receptors (Canet et al., 2019a) may be a viable option for slowing the disease progression, regardless of when treatment is applied. The need for mitigating treatments is evident in light of the failure of amyloid-reducing therapeutics to delay cognitive decline (Cummings et al., 2019; Oxford et al., 2020; Panza et al., 2019). That is, while glucocorticoid-targeting treatments are unlikely to restore cognitive function, they may be able to reduce the rate of hippocampal structural decline, prolonging the maintenance of hippocampal function, even after disease initiation.
Alternatively, if HPA-axis dysfunction occurs much earlier in the disease process (i.e., before decrements in hippocampal-dependent functioning are generally observed) and early hypercortisolemia is a driver of other pathologies (such as Aβ accumulation, as discussed in the Cortisol and Amyloid Interaction section), therapeutic interventions targeting glucocorticoids may be able to delay or prevent the onset of the disease. The results of this meta-analysis, in particular the finding that basal cortisol elevations are far from a consistent or ubiquitous observation across studies as is often assumed, suggest that efforts to identify the timing and moderating factors of hypercortisolemia in AD could pave the way for effective treatments.

**Covariates**

*Fluid*

The biological fluid collected and assayed for cortisol was not a statistically significant predictor of between-study heterogeneity (see Figure 10). This was somewhat surprising because although all three biological fluids are valid measures of cortisol, there are differences in how they are typically utilized in research that lead to the assessment of slightly different aspects of HPA axis functioning (Spencer & Deak, 2017).

Under basal conditions, a distinct circadian pattern of cortisol concentrations is produced by differences in the length and frequency of cortisol secretory pulses. The resultant pattern is characterized by a morning peak just after awakening followed by a decline throughout the day and a quiescent period during sleep (Bliss et al., 1953; Weitzman et al., 1971). Cortisol concentrations observed during the morning peak are approximately four times those observed during the circadian trough (Spencer & Deak, 2017) making direct comparison of levels across different times of day uninformative.
As shown in Figure 10, blood was the most frequently utilized collection fluid. Due to the secretion of cortisol into blood from the adrenal glands collection and assay of cortisol from blood samples (plasma or serum) are considered the most direct measure of circulating cortisol. However, there are two relevant limitations associated with the use of blood sampling: the requirement of collection in a clinical setting and the assay of total cortisol levels. The requirement of a clinical setting for blood collection impacts the obtained samples in two ways. First, the typical protocol includes just one collection because participants often do not remain in the clinical setting for more than a few hours. This single collection reflects the circulating cortisol concentration at just the minute of collection. This introduces random between-subject variation due to the pulsatile nature of cortisol secretion. That is, observed concentrations will differ as a function of both true differences in basal cortisol secretion and as a function of the time since the last cortisol secretion burst, as well as the duration of that secretory burst. Because information regarding the latter relationship cannot be determined from a single collection, it is simply accepted as noise.

The second impact of the requirement of a clinical setting for collection is the tendency to align collections to a set time-of-day (e.g., 09:00) without regard for participant waking time. Because the circadian rhythm of cortisol is determined by sleeping and waking patterns rather than specific, set times of day, individual differences in waking time (and typical sleep-wake times) manifest as between-subject differences in cortisol in time-of-day based collections. This is particularly apparent for morning collections due to the sharp rise and subsequent fall associated with waking, whereas levels throughout the remaining waking day are generally more stable.
In addition to noise introduced by the typical sampling protocols, cortisol levels assayed from blood samples are most commonly total, rather than free cortisol levels (Gatti et al., 2009; Spencer & Deak, 2017). Most circulating cortisol is bound to cortisol binding globulin, with a smaller fraction bound to albumin. This bound cortisol is unable to directly cross the blood-brain barrier and thus cannot bind to neural receptors such as those in the hippocampus (Mendel, 1989). Therefore, although it is the much smaller unbound or free portion of cortisol that is of primary interest in the context of the GCH (Perogamvros et al., 2012), levels assayed from blood typically reflect mostly total cortisol. While the free and total cortisol levels are often highly correlated, this correlation declines at higher levels of circulating cortisol (Aardal & Holm, 1995).

The use of salivary collection addresses several of the problems associated with cortisol sampling for blood. First, salivary sampling eliminates the need for a clinical setting for collection and typical protocols involve providing participants (or caretakers) with collection devices (generally sterile tubes and cotton swabs) to be used at home. This approach allows for collection schedules relevant to participants’ sleep-wake schedules (e.g., samples obtained at waking, thirty minutes after waking, and one hour after waking), reducing between-subject variation introduced by fixed collection times (e.g., 09:00). Salivary cortisol sampling protocols also typically include multiple collections across the day, likely due to the at-home rather than clinical setting of enabled by salivary collections. This is evident in the current analysis, as all effects calculated using indices derived salivary collections included at least two unique collection times.

In addition to allowing for collection outside of a clinical setting, transport from the blood to saliva requires passage through several lipophilic layers. Therefore, much
like the blood-brain-barrier, the passage of bound cortisol is largely prohibited. Thus, assays of salivary collections are reflections of free or unbound cortisol, rather than total circulating cortisol (Gröschl, 2008; Hellhammer et al., 2009).

Salivary sampling is not without limitations, however. First, like cortisol assayed from blood samples, salivary cortisol reflects cortisol concentrations only at the time of collection. Therefore, between-subject differences in the timing of pulsatile secretions in relation to sampling still introduce noise. Second, the unsupervised nature of collection common in protocols can introduce participant non-compliance, particularly for waking measures (Almela et al., 2012).

Finally, urinary cortisol sampling, while typically requiring collection in a clinical setting, provides a measure of free, rather than total cortisol and overcomes the limitation of momentary sampling observed for blood and saliva sampling by providing an integrated measure of cortisol exposure over a period of hours (Remer et al., 2008). For all effects derived from urinary samples included here, the protocol required 24-h sampling thereby eliminating the noise introduced by between-subject differences in the timing of the circadian rhythm in relation to sample collection. Despite these advantages, concerns of high cross-reactivity of assay techniques and invalid assessment of participant compliance have been voiced in relation to urinary cortisol measurements (Fenske, 2006; John et al., 2016; Murphy, B. E. P., 2002).

Given the impact of the circadian rhythm and pulsatile secretions on individual plasma and salivary samples, it was predicted that effects determined from urinary cortisol measurements would be the largest and most consistent, followed by those effects determined from multiple collections (typical of salivary protocols, those also included in
several blood sampling protocols, notably Hartmann et al. (1997) and Davis, K. L. et al. (1986) then finally effects derived from single-moment collections. This finding is even more surprising given that the majority of single-moment plasma measures were obtained in the morning, when small differences in timing would seemingly introduce the most amount of noise. However, these findings are in line with the meta-analytic findings of Murri et al. (2014). This group reported that basal cortisol levels in depressed patients were an average of 0.89, 0.83 and 1.39 SDs higher than older control participants when measured in the morning, afternoon, or at night respectively, but just 0.51 SDs higher when basal cortisol was measured from “continuous” (i.e., encompassing a 12-24 hour period) collections.

Although counter to predictions, the effect size invariance in relation to fluid demonstrates the robustness of the difference in cortisol levels between AD and NCC older adults, as even in situations where the measurement of cortisol is particularly noisy, the presence of AD was associated with increased cortisol levels. Further, the effect size invariance to fluid indirectly supports the assumption made on the basis of findings from Davis, B. M. et al. (1985) and Hartmann et al. (1997) that differences in cortisol level between AD and NCC older adults do not systematically differ as a function of time of day.

It bears mentioning that no effects calculated from CSF were represented in any analyses here. Although CSF may provide the most accurate estimate of neural exposure to free cortisol, CSF is difficult to obtain and is therefore rarely sampled in healthy

---

3 This article was not directly included in later analyses but the information contained is included via Davis, K. L. et al. (1986), see Effect 12.
participants. One notable exception is the sample pair from Effect 47 reflecting participants from the German Dementia Competence Network (Kornhuber et al., 2009). The protocol for this cohort study included the collection of both blood and CSF for cortisol assay. The effect size calculated from the blood collection was selected for inclusion in this meta-analysis to maintain consistency with the other included effects, none of which were determined from CSF. However, in this sample pair, the AD sample had significantly \( (d = 0.85, 95\% \text{ CI } = (0.46, 1.24)) \) higher CSF cortisol levels \( (M = 0.56, \text{SD} = 0.39) \) than the NCC sample \( (M = 0.25, \text{SD} = 0.25) \). It is notable that this effect is more than double the average observed effect size from Analysis Two and differs from the effect observed in the same participants using blood sampling \( (d = -0.18, 95\% \text{ CI } = -0.55 - 0.20) \) (see Ouanes & Popp, 2019 for discussion).

Despite the strong effect observed by Popp et al. (2015), cortisol assayed from CSF does not appear to be immune to the heterogeneity observed in measurements from other fluids. Peña-Bautista et al. (2019) compared cortisol assayed from plasma, urine, CSF, and saliva obtained from participants between 08:00 and 10:00. This group reported that AD was associated with higher plasma and higher urinary cortisol levels but there was no association with salivary or CSF levels. Unfortunately, none of the reported effects could be included in the present analysis due to the use of a combined normal cognition and non-AD dementia (i.e., fronto-temporal, vascular and Lewy body dementia or psychiatric disorders such as schizophrenia, depressive disorder and bipolar disorder) control sample. Similarly, Laske et al. (2009) reported that AD patients had significantly elevated morning serum cortisol levels compared to healthy controls (Effect 25),
however, serum cortisol measures did not significantly correlate with CSF cortisol measures taken directly after serum collection in the AD sample 4.

Finally, although implied by the use of the term Fluid to describe the biological specimens used for cortisol assay, cortisol levels assayed from hair samples were not explicitly excluded from this analysis. However, no studies met the remaining eligibility criteria. For example, de la Rubia Ortí et al. (2019) collected and assayed hair cortisol samples from 49 participants with mild, moderate, or severe AD. Significant between-group differences were reported, however no cognitively normal control group was included. While still in earlier stages of validation than blood, saliva or urine cortisol measures (Burnard et al., 2017; Short et al., 2016), hair cortisol levels are hypothesized to reflect cortisol exposure over a period of months (rather than minutes or days). Elevated hair cortisol levels have been reported in Cushing syndrome patients (Hodes et al., 2017; Manenschijn et al., 2012; Thomson et al., 2010), dementia caregivers (Stalder et al., 2014), and first-episode depressive patients (Wei et al., 2015), suggesting comparison of hair cortisol from cognitively normal and demented patients may be a viable avenue of future research (Liu & Doan, 2019; Wright et al., 2015).

**Age**

Age was examined as a potential influence on the SMD of cortisol levels on the basis of research showing that cortisol levels increase with advancing age, even in cognitively normal older adults (Deuschle et al., 1997; Nicolson et al., 1997; Purnell et al., 2004; Rueggeberg et al., 2012). Therefore, it was anticipated that sample pairs which

4 No CSF cortisol measures were reported for the healthy control group, therefore this effect was ineligible for inclusion in the present analysis.
included an older AD sample than NCC sample would have larger positive SMDs than sample pairs with matched ages or an older NCC sample. As shown in Table 3 and Figure 11, although most sample pairs included a numerically older AD sample than NCC sample and there was variation in the age difference between groups, no relationship between the age difference and cortisol difference was observed.

Figure 13 depicts three hypothetical relationships between cortisol levels and age in AD and NCC participants. All three panels depict a small positive linear relationship between age and cortisol levels for cognitively normal older adults. Panel A shows an exponential relationship between cortisol levels and age for older adults with AD indicative of the GCH whereby initial cortisol elevations invite hippocampal damage, reducing inhibition of further basal increases.
Figure 13. Illustration of three hypothetical relationships between cortisol and age for adults with AD and NCCs. All three panels depict a small positive linear relationship between age and cortisol for the NCC group, illustrating the assumption that even those who do not show cognitive deficits have increased cortisol levels with age. Panel A depicts an exponential relationship between cortisol levels and age in the AD patients. Panel B depicts a linear relationship between cortisol and age for the AD patients, with a steeper slope than the NCC group. Figure C depicts a linear relationship such that AD patients and NCC patients have parallel slopes.
The results from Analysis Four do not provide support for the scenario depicted in Panel A. Drawing participants from such a scenario would be expected to result in significant weighted age and difference terms in the context of Analysis Four. That is, in the case of well age-matched sample pairs with a relatively low weighted mean age, the expected SMD of cortisol levels would be very small, compared to sample pairs with higher mean weighted ages, producing a positive regression coefficient for the weighted mean age term. Similarly, sample pairs with an older AD sample than NCC sample (and therefore, a positive age difference term) would also be expected to have larger SMDs in cortisol level. The opposite would be anticipated in sample pairs with older NCC than AD samples, thereby again producing a positive regression coefficient for the difference in age term. Although not tested in any analyses, if the AD samples included in this analysis were drawn from such an exponential distribution, a positive interaction between weighted mean age and difference in age would also be anticipated.

Panel B depicts a second hypothetical relationship between cortisol levels and age for AD patients. This relationship suggests that both NCC and AD older adults experience linear increases in cortisol levels, however, AD older adults experience a steeper increase. While the present results do not lend extensive support to this scenario, it may be that the increase in the cortisol slope in AD individuals is subtle, and therefore could not be statistically detected in the samples included here, which included fairly restricted age ranges (see panel A of Figure 11).

Finally, Panel C reflects a scenario in which older adult participants have identical change slopes regardless of cognitive status but individuals with AD have a higher cortisol intercept than NCC individuals. The results of the present analyses most
closely model the scenario depicted in this panel. That is, regardless of the weighted age of the sample pair, or the difference in the mean age of the AD and NCC samples, the SMD in cortisol between the two groups remains the same, as suggested by the insignificant regression weights. Notably, while Panel C reflects a small positive linear slope for both groups, the results of this analysis do not provide any information regarding the shared slope.

Additionally, the results of Analysis Four are not informative of the time course or process by which AD patients have a higher cortisol intercept. That is, it is possible that eventual AD patients have elevated cortisol levels which are evident years prior to diagnosis, as suggested by Ennis et al. (2017). Such a scenario would not be detected here because cortisol measures were only measured in AD patients after diagnosis (or, at a point when cognitive deficits are present). Assuming that cortisol elevations are present well in advance of observed clinical impairment, the number of future AD participants (i.e., those who are cognitively normal at the time of cortisol measurement despite being on the altered trajectory) included in the NCC samples is a potential source of between-effect heterogeneity which was unmeasured in all of the present analyses.

Sex

As shown in Table 3, Analysis Four the proportion of female participants included in the sample pairs did not impact the size of the SMD in cortisol levels. Reports of sex differences in AD are common (see Laws et al., 2018) including accelerated hippocampal atrophy (Ardekani et al., 2016) and cortisol elevations (Leblhuber et al., 1993) in female compared to male AD patients. However, these results are not unanimous (Intebi et al., 2002). While there are a number of candidate mechanisms that may lead to an
expectation of sex differences in cortisol effects, including differences in vulnerability to chronic stress (Yan et al., 2018), the stress response (Thomas, N. et al., 2019), energy and amyloid metabolism (Zhao, L. et al., 2016) or sex hormones (Vest & Pike, 2013; Zhao, L. et al., 2016), speculation is not warranted based on the null finding.

**MMSE Score**

The MMSE is a short, easily administered neuropsychological test of global cognitive functioning which shows high diagnostic accuracy for dementia (see Larner, 2019) and high test-retest reliability (Pangman et al., 2000). It was used here as a surrogate measure for the clinical severity of AD due to its observed agreement with Clinical Dementia Ratings (Perneczky et al., 2006), wide range of possible scores (0 – 30), and frequent use in the literature (reported for 39 of 63 effects here, see Figure 12).

Somewhat surprisingly, given previous reports of negative relationships between plasma (Huang et al., 2009; Zvěřová et al., 2013) and even CSF (Gil-Bea et al., 2010) cortisol levels and MMSE scores in AD participants, the MMSE score of the AD sample did not predict a significant portion of the between-study heterogeneity. Assuming that MMSE score is an adequate metric of dementia severity (see Flicker, 2017 for discussion), the null association between basal cortisol levels and MMSE score is in line with earlier reports of a lack of concordance between the level of biological measures of AD and the severity of the Alzheimer’s clinical syndrome (Gottfries et al., 1994; Gurevitch et al., 1989).

There is evidence that education (and associated but quantifiably difficult factors such as IQ and occupational attainment) often moderate the biomarker-clinical severity relationship (Ewers, 2020; Meng & D’Arcy, 2012). This moderation effect may be
particularly influential here, because even in non-demented populations, education levels impact MMSE performance (Crum et al., 1993). Therefore, although education levels were rarely reported in the primary studies, it is speculated that education influenced the meta-analytic findings here. Examination of several individual studies provides some support for this speculation. Specifically, both the BLSA and ADNI samples included here are highly educated (16.4 years and 15.6 years (Wirth et al., 2019) of education respectively) and all three associated AD samples have mean MMSE scores near or above the recommended cut-point for dementia (26.7, 24.6, and 23.6 for BLSA females, BLSA males and the ADNI sample respectively) (Mitchell, 2013), despite meeting the NINCDS-ADRDA criteria for probable AD. In stark contrast, Umegaki et al. (2000), although not reporting education levels, included an AD sample with a mean MMSE score of 7.96 and a NCC sample with a mean MMSE score of 24.5. Given this contrast it seems unreasonable to assume that between-study differences in MMSE score strictly reflect differences in AD biological progression.

**Unexamined Influences**

As shown in Table 3, even after consideration of the influences of biological fluid, age, sex, and MMSE score, significant residual between-study heterogeneity remained. While a plethora of factors are associated with cortisol levels and/or AD, this discussion focuses on a small subset of those factors hypothesized to systematically influence the difference in observed cortisol levels between AD and NCC samples.

**Disease Severity and Duration**

As previously discussed, every modern definition of AD has in some way distinguished between the biological etiology of the *disease* and the manifestation of
cognitive *clinical* symptoms (Jack et al., 2018; McKhann et al., 1984). Because the diagnostic criteria used to distinguish the AD and NCC samples included here (and throughout the AD literature) focus on the clinical symptoms, indicators of disease severity aside from cortisol (such as measures of Aβ or neural structural markers (Counts et al., 2017)) were rarely measured and even less frequently reported in the primary studies. Similarly, estimates of the duration of disease were rarely and inconsistently reported, prohibiting meaningful systematic examination here.

**Depression**

History of clinical depression or depressive symptoms has consistently been associated with both hypercortisolism (Murri et al., 2014; Stetler & Miller, 2011) and increased risk of AD (Barnes et al., 2012; Cherbuin et al., 2015; Murri et al., 2014). Further the co-morbidity of dementia and depression has been associated with more severe HPA-axis dysregulation than those observed in either condition alone (Barca et al., 2019; O'Brien et al., 1996). However, like measures of disease severity or duration, depressive status or depressive symptoms were not reported frequently or consistently enough across studies to be meaningfully examined at the meta-analytic level.

**Other Indicators of HPA-Axis Function**

In addition to basal cortisol levels, a number validated indicators of HPA-axis functioning exist in the literature, such as the CAR (Clow et al., 2004), circadian rhythm flattening (Spiga et al., 2014), effectiveness of dexamethasone suppression (Golden et al., 2011), or cortisol reactivity (Kirschbaum et al., 1993). These indicators, while of interest even in the context of AD, have substantially different conceptual interpretations and therefore are beyond the scope of this meta-analysis.
Diagnostic Criteria

As described in the Cognitive Status Grouping section, the author designation of dementia was accepted, regardless of the diagnostic criteria used. The first commonly utilized clinical diagnostic criteria for AD were those put forth by the 1984 NINCDS-ADRDA workgroup (McKhann et al., 1984). Several years later the DSM-III-R (American Psychiatric Association, 1987) and ICD-10 (World Health Organization, 1992) introduced AD-specific categorizations, evolving from earlier more broadly defined categories of “Organic Brain Syndrome” (American Psychiatric Association, 1952, 1968; World Health Organization, 1974, 1978) and “Primary Degenerative Dementia” (American Psychiatric Association, 1980).

The AD diagnostic criteria widely differ with regard to the pattern and severity of the clinical profile necessary to warrant a positive diagnosis (George et al., 2011; Kane & Thomas, 2017; Lauter et al., 1990; Storey et al., 2001), the resulting estimated prevalence of dementia, (Erkinjuntti et al., 1997) and association with histopathological patterns (Gaugler et al., 2013; Jobst et al., 1997). This variance coupled with the frequent use of non-standard or unreported methods for AD diagnosis in the primary studies prohibits the inclusion of diagnostic method as a meaningful or interpretable mediating variable at the meta-analytic level. Likewise, the comparison of basal levels between NCC or AD samples with MCI samples (see Johar et al., 2015; Popp et al., 2009; Popp et al., 2015; Swanwick et al., 1998), or within-AD sample comparisons by severity (see James et al., 2019; Pomara et al., 1984) are of interest but these classifications are not assessed or reported with enough consistency to meaningfully examine at the meta-analytic level.
Limitations

In addition to the usual limitations of meta-analyses of cross-sectional literature, two aspects of this literature specifically warrant consideration. First, cortisol, like many biological measures, is commonly positively skewed. The presence of positive skew in the primary data may have introduced slight inaccuracy when estimating the mean and SD from other measures of central tendency and spread, as the estimation procedures may assume data is normally distributed (Luo et al., 2018; Wan et al., 2014). Additionally, in a meta-analytic simulation study, Sun and Cheung (2020) showed that analyses conducted using skewed primary distributions had an increased Type I error rate. Given the size of the effect and very small p-value (< 0.0001) associated with Analysis Two, it is unlikely that the main finding here is a false positive.

Second, it is assumed that the observed effect is driven by small to moderate cortisol elevations in the whole AD sample. However, it is possible that larger elevations in just a small subset of patients is responsible for the observed result. Although this cannot be directly addressed at the meta-analytic level, the latter scenario could be statistically problematic because it would likely create systematic differences in the primary sample variances, violating the homogeneity of variances assumption of Hedges g (Hedges, 1981). Recently, Aoki (2020) introduced $e$ (and corresponding package es.dif), a measure of the SMD without the assumption of homogeneity of variances. As a post-hoc examination, Analyses Two through Five were conducted using $e$. No significant or interpretative differences from the reported results emerged.
Summary and Conclusions

This study reports the only meta-analytic examination of differences in basal cortisol levels between older adults with AD and NCC adults. The unadjusted model showed that unstimulated cortisol levels are almost one half of a SD higher in AD than in non-demented controls. These results support the hypothesis that HPA-axis dysregulation is involved in the etiology of AD. However, it is still ambiguous when dysregulation begins in the disease process and how basal cortisol elevations impact the disease and associated clinical syndrome.

The combined results presented here show that significant between-study variance exists which cannot be accounted for by differences in the biological fluid used for cortisol measurement, sample mean ages, the proportion of female participants or a simple measure of clinical severity. Further research is warranted to examine the time course and trajectory of HPA-axis dysregulation in AD and to quantify the influence of factors which contribute systematically to between-study effect size differences.
CHAPTER 3: CORTISOL AND AMYLOID DEPOSITION RISK

The preceding meta-analytic results suggest that basal cortisol elevations are common in the Alzheimer’s clinical syndrome, however, as stated in the Limitations section, these findings provided no information about the time course of the cortisol involvement in the biological progression of AD. Therefore, the present analysis aims to determine whether cortisol levels obtained prior to PET imaging evidence of Aβ accumulation are informative of accumulation risk.

Prediction of Aβ Accumulation

As discussed in the In Vivo Imaging section, the development of a radiolabeled compounds with an affinity for Aβ for use with PET imaging revolutionized the study of AD by allowing for quantification of Aβ in-vivo (see Young et al., 2020). Aβ accumulation is generally hypothesized to precede and contribute to the decline of cognitive functioning (Hardy & Higgins, 1992). On the basis of the failure of anti-amyloid drugs to restore or maintain cognitive functioning (Cummings et al., 2019; Oxford et al., 2020; Panza et al., 2019) researchers have hypothesized that the prevention of AD pathology, in particular Aβ accumulation, is the only viable route for treatment (Viña & Sanz-Ros, 2018).

In an effort to identify viable therapeutic targets, researchers have examined the relationship between several biological and lifestyle factors and Aβ pathology, both in vivo (e.g., Resnick et al., 2015) and post-mortem (e.g., Ashby et al., 2016). One major limitation to this published research, due in part to the predominant availability of case-controlled studies of AD (Song & Chung, 2010), is that even for potentially time-varying
measures of interest (such as cholesterol or blood pressure), measures were obtained concurrently with PET imaging from samples containing a mix of Aβ+ and Aβ- participants (Gomez et al., 2018; Köbe et al., 2020; Rodrigue et al., 2013; Toledo et al., 2012). Therefore, these designs allow for conclusions regarding group differences but cannot address whether differences in these biomarkers are present prior to the initiation of Aβ accumulation.

**Current Investigation**

The present investigation is uniquely positioned in this literature because it examines the influence of a biomarker level obtained from all participants prior to any evidence of Aβ accumulation. Through the application of a Cox regression analysis (Cox, 1972) to data collected from an extensive longitudinal study, this investigation was able to address whether individual differences in basal cortisol levels, or changes in cortisol level over a period of years, increase the risk of subsequent Aβ accumulation.

Given the specific focus on the time frame prior to Aβ accumulation, cortisol is a logical choice for examination for several reasons. First, as previously discussed, cortisol elevations have been associated with increased AD risk (Doecke et al., 2012; Ennis et al., 2017) suggesting that basal cortisol elevations are evident in advance of overt cognitive symptoms. Similarly, cortisol has been proposed as the mediator between depression, a condition characterized by extended basal cortisol elevations (Murri et al., 2014; Stetler & Miller, 2011), and the epidemiologically observed increased risk of AD in this population (Canet et al., 2018; Sotiropoulos et al., 2008). Further, evidence suggests that this mediating effect is driven by increases in Aβ accumulation in depressed individuals.
(Chung et al., 2015; Harrington et al., 2015; Krell-Roesch et al., 2018; Krell-Roesch et al., 2019; Lavretsky et al., 2009; Wu, K.-Y. et al., 2014).

Second, in vitro and rodent studies have demonstrated that glucocorticoids act in a dose-dependent manner to increase Aβ secretion (Dong et al., 2004; Green et al., 2006; Wang, Y. et al., 2011). This evidence of a direct interaction between cortisol and Aβ further suggests that cortisol elevations may serve to trigger Aβ accumulation. Finally, Toledo et al. (2012) reported a cross-sectional association between morning plasma cortisol levels and higher Aβ.

The present study extends earlier work by investigating whether cortisol elevations prior to initial Aβ accumulation are predictive of future Aβ accumulation. This investigation leverages both the longevity of the BLSA protocols and the unique availability of 24-h urinary cortisol measurements obtained years to decades prior to PIB imaging. The present analysis of 799 urinary cortisol samples and 429 PIB scans from 145 participants for up to 35 years reflects the largest longitudinal evaluation of the cortisol-Aβ hypothesis yet undertaken. It is hypothesized that elevated 24-h basal cortisol levels and increasing cortisol trajectories will be associated with increased risk of Aβ accumulation.
Method

Participants

The participants included in this study are a subset of individuals from the BLSA (Shock et al., 1984) included in both the neuroimaging study (BLSA-NI) (Resnick et al., 2000) and cortisol assay subset. The BLSA consists of volunteers who return to the NIA in Baltimore biannually where they receive a number of behavioral and medical assessments and provide specimen samples, including a 24-hr urine sample which is subsequently frozen and stored in the BLSA specimen bank. Six thousand and sixty-three of these urine samples contributed by 1,865 individuals were selected for cortisol assay on the basis of concurrent cognitive testing. The BLSA-NI was initiated in 1994 and includes individuals who underwent annual MRI and PET scans (Resnick et al., 2000). One hundred and forty-five of the BLSA-NI participants also had available cortisol measures.

Cortisol

The cortisol measures used in the present study were obtained using the 24-hr urine samples provided as part of their biennial assessments. Participants were instructed to void upon waking at approximately 8 am and discard this specimen. Following initial voiding, all urine including the final specimen voided at the end of the 24-hour collection was collected by the subject in containers provided to them. Cortisol levels have been demonstrated to remain stable following 24-hr of storage at room temperature without preservative (Gouarne et al., 2004). Subsequently, multiple 20 mL samples of urine were aliquoted from the total pool and stored at -80° C in the BLSA specimen bank. Cortisol and creatinine are stable in urine, even when frozen for long durations (Miki & Sudo, 1998; Soliman et al., 1986).

Assays were conducted by Esoterix Incorporated. UFC was measured by liquid chromatography with mass spectrometry. Interassay coefficients of variation for mean
values of 0.05, 2.84, 5.53 and 9.41 ug/dl were 13.4%, 6.6%, 4.4% and 5.3% respectively. Creatinine concentrations were determined by Esoterix Inc. using the Synchron LX System (Beckman Coulter, 1998). Interassay coefficients of variation for creatinine mean values of 66.6 ng/mL and 145.8 ng/mL were 1.5% and 1.1% respectively. The ratio of UFC to creatinine (μg:g) (hereafter referred to as UCR) was calculated for each specimen, a standard practice consistent with prior studies (Kacsoh, 2000, pp. 115-119; Seeman et al., 1997). Correction of urinary cortisol concentration by creatinine excretion accounts for individual differences in urine dilution (due to individual variation in daily fluid intake) and corrects for age-related differences in kidney filtration rate. (Alessio et al., 1985; Driver & McAlevy, 1980).

**PIB**

All PIB data was collected through a standardized protocol utilized by the BLSA described in Resnick et al. (2015). Dynamic PIB PET was conducted using a GE Advance scanner in 3D mode. Participants were fitted with a thermoplastic head-mask, then administered a bolus IV injection of PIB. Scanning began immediately and proceeded for 70 minutes for a total of 33 frames. Dynamic images were reconstructed such that the full width at half maximum at the center of field of view was 4.5 mm. Automated anatomical labeling was achieved by affinely registering the Statistical Parametric Mapping PET template onto the 20-minute mean PET-PIB image for each scan (Tzourio-Mazoyer et al., 2002). Quantification of DVR was performed as described by Zhou et al. (2007) as the simplified reference tissue model using the cerebellum as the reference region. Mean cortical DVR (cDVR), a global DVR measure, was calculated as the average of DVR values in superior, middle, and inferior frontal and orbitofrontal,
superior parietal, supramarginal, and angular gyrus regions, precuneus, superior, middle, and inferior occipital, superior, middle, and inferior temporal, anterior, middle, and posterior cingulate regions.

A cDVR value of exactly 1 indicates that the amount of PIB binding in the cortex is equal to the binding observed in the reference region (i.e., the cerebellum), while higher values indicate that more PIB is bound in the cortex than reference region. For the purposes of this study, scans were designated as showing evidence of accumulation (i.e., PIB+) if the cDVR was greater than or equal to 1.062 (Lopresti et al., 2005; Sojkova et al., 2011).

Figure 14 shows the timeline of UCR and PIB collections in the BLSA. This information is presented in Figure 15 separated by participant and aligned to the PIB baseline measure. See Table 4 for a clarification of terminology.
Figure 14. The count, by year and type of collection, of the PIB and UCR collections only for those participants (N = 145) with at least one UCR measure and one PIB measure. A total of 799 UCR measures and 429 PIB measures were collected from these participants. Each bar represents a calendar year. PIB and UCR were not necessarily measured at the same visit within-participant. Therefore, overlapping bars indicate that in the entire sample of participants some count of both measures were collected in the same calendar year, but these overlapping measures were not necessarily collected from the same participant(s).
Figure 15. The timing of UCR and PIB measurements in relation to the baseline PIB measurement for all participants with at least one UCR measure and at least one PIB measure. Points reflect the amount of time (rounded to the nearest year) between the observed measure (UCR or PIB) and the PIB baseline. All visits per participant appear in a single row connected by the gray dotted line.
Table 4: Specification of Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIB+ scan</td>
<td>A PET scan for which the calculated cDVR of PIB was $\geq 1.062$</td>
</tr>
<tr>
<td>PIB- scan</td>
<td>A PET scan for which the calculated cDVR of PIB was $\leq 1.062$</td>
</tr>
<tr>
<td>PIB baseline</td>
<td>The first PIB scan obtained for each participant by the BLSA (regardless of scan or participant PIB+/PIB- status)</td>
</tr>
<tr>
<td>PIB+ participant</td>
<td>A participant with one or more PIB+ scans obtained by the BLSA</td>
</tr>
<tr>
<td>PIB- participant</td>
<td>A participant with no PIB+ scans obtained by the BLSA</td>
</tr>
<tr>
<td>PIB intercept</td>
<td>The first PIB+ scan obtained (determined only for PIB+ participants)</td>
</tr>
<tr>
<td>UCR baseline</td>
<td>The first measurement of UCR by the BLSA for each participant (regardless of PIB+/PIB- scan or participant status)</td>
</tr>
<tr>
<td>Eligible UCR</td>
<td>A measurement of UCR eligible for inclusion in the calculation of the UCR Summary measures. To be considered eligible for inclusion, the UCR measure had to be collected prior to the event time.</td>
</tr>
<tr>
<td>Last eligible UCR</td>
<td>The eligible UCR measure for each participant occurring latest in time</td>
</tr>
<tr>
<td>Time-to-event</td>
<td>The amount of time between the last eligible UCR measure and the event time for each participant</td>
</tr>
<tr>
<td>Event time</td>
<td>For PIB+ participants, the time at which the participant was predicted to have a cDVR of exactly 1.062. For PIB- participants, the time of the last obtained PIB scan (inherently, a PIB- scan) by the BLSA</td>
</tr>
</tbody>
</table>
Analytic Approach

This study used Cox regression models to investigate whether UCR measures collected prior to PET imaging evidence of Aβ accumulation are associated with the subsequent risk of initiating Aβ accumulation in older adulthood.

Estimation of Event Time

For the PIB+ participants, the event time was defined as the time at which it is estimated that the participant had cerebral Aβ accumulation equivalent to the level of accumulation that, if scanned, would result in a PIB cDVR value (as calculated above) of exactly 1.062. Event time for all PIB+ participants was estimated using a linear model. First, a linear mixed effects model was applied to the cDVR values from all PIB+ scans grouped by participant. The cDVR from the first PIB+ scan available for each PIB+ participant was used as the intercept, designated as Time 0. The average of the within-participant slopes was then used to calculate the time between Time 0 (i.e., the PIB intercept) and a predicted cDVR of 1.062.

For participants that did not show any evidence of PIB accumulation in any of the PIB scans obtained by the BLSA (i.e., the PIB- participants), the event time was the time of the last recorded PIB scan for each participant. The last PIB scan for the PIB- participants was inherently PIB- (i.e., have a cDVR value less than 1.062).

Cortisol Summary Measures

In order to assess the influence time invariant longitudinal UCR Summary measures, only participants with two or more eligible UCR measures were included in the analysis. For these participants the UCR Mean, defined as the mathematical average of all eligible UCR values, the UCR slope, defined as the slope of the least squares regression
line for all eligible UCR values, and the UCR Closest, defined as the last eligible UCR measure were determined.

**Cox Regression Models**

Cox regression models were fit using the Survival package (Therneau, 2020; Therneau & Grambsch, 2000) in R version 4.0.1 (R Core Team, 2020), using the exact method for event time ties. First, an unadjusted model for each of the three UCR Summary measures (i.e., UCR Mean, UCR Slope, and UCR Closest), followed by an adjusted model for each of the three measures with an added age term.
Results

Demographics and relevant group comparisons are shown in Table 5. Participants with only one UCR measure prior to event time had significantly fewer UCR measures, spanning a significantly shorter number of years, compared to those participants with two UCR measures prior to event time. Of the 115 participants with two or more UCR measures collected pre-event, the PIB- group had significantly more UCR measures spanning a significantly longer follow-up time than the PIB+ group. The two groups had significantly different cortisol slopes, such that the PIB+ group had an average positive (i.e., increasing UCR levels over time) slope whereas the PIB- group had an average negative slope (i.e., decreasing UCR levels over time). Regarding other demographic variables, the two groups only differed on the use of statin medication, such that 20% of the PIB+ group reported using statins at any of the UCR measurement occasions, while 60% of the PIB- participants reported using statins.
Table 5: Demographic Comparisons

<table>
<thead>
<tr>
<th>Has 2+ UCR Pre-Event</th>
<th>PIB Status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
</tr>
<tr>
<td>Age: BL</td>
<td>60.55 (9.41)</td>
</tr>
<tr>
<td>Age: LV</td>
<td>68.66 (14.44)</td>
</tr>
<tr>
<td>Age: Mean</td>
<td>64.98 (11.78)</td>
</tr>
<tr>
<td>UCR: N</td>
<td>4.57 (3.40)</td>
</tr>
<tr>
<td>UCR: Years</td>
<td>8.12 (7.61)</td>
</tr>
<tr>
<td>UCR: Mean</td>
<td>16.80 (9.28)</td>
</tr>
<tr>
<td>Education</td>
<td>-0.27 (0.89)</td>
</tr>
<tr>
<td>MMSE</td>
<td>16.67 (2.89)</td>
</tr>
<tr>
<td>Female</td>
<td>13 (43.3)</td>
</tr>
<tr>
<td>White</td>
<td>21 (70.0)</td>
</tr>
<tr>
<td>APOE e4:</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>14 (46.7)</td>
</tr>
<tr>
<td>Yes</td>
<td>15 (50.0)</td>
</tr>
<tr>
<td>Missing</td>
<td>1 (3.3)</td>
</tr>
<tr>
<td>BMI</td>
<td>27.11 (3.97)</td>
</tr>
<tr>
<td>HDL</td>
<td>52.38 (11.97)</td>
</tr>
<tr>
<td>LDL</td>
<td>112.78 (24.76)</td>
</tr>
<tr>
<td>CESD</td>
<td>4.65 (4.00)</td>
</tr>
<tr>
<td>Smoker</td>
<td>13 (43.3)</td>
</tr>
<tr>
<td>Diuretic</td>
<td>9 (30.0)</td>
</tr>
<tr>
<td>Statin</td>
<td>16 (53.3)</td>
</tr>
<tr>
<td>Steroid</td>
<td>10 (33.3)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>12 (40.0)</td>
</tr>
<tr>
<td>HTN</td>
<td>6 (20.0)</td>
</tr>
</tbody>
</table>

Note. Var: Variability. Values for all continuous variables are represented as mean (SD) and categorical variables are count (percentage). The Age LV reflects the age at the last eligible UCR measure. Values reflect information obtained at UCR collection visits and do not include time-varying demographic values relevant to the PIB scan visits. The mean of each continuous variable was calculated using all available measures matching inclusion criteria. All categorical variables represent the number of participants in the relevant group that were classified as ever yes for the relevant demographic measure, only during any of those visits which were included in mean estimations.

\( a p < 0.05, b p < 0.01, c p < 0.001 \). All p-values correspond to t-test or chi-square test outcomes comparing the values in the Yes/PIB+ column to the values in the No/PIB- column directly to the left.
**Estimation of Event Time**

The results of the mixed effects linear model showed that among the 154 cDVR values greater than 1.062 collected from 55 participants with at least one PIB+ measure, cDVR increased at a rate of 0.023 units \((t = 13.02)\) per year. Further, there was very little variance in the slopes across participants (variance = 0.000059). The intercept \((b_0)\), reflecting the average cDVR at the first PIB+ visit, was significantly different than zero \((b_0 = 1.25, t = 52.72)\).

As depicted in Panel B of Figure 16, the event time (as years to PIB+ intercept, \(y\)) for each PIB accumulator \((i)\) was calculated using the intercept cDVR \((b_0)\) as:

\[
y_i = \frac{b_{0i} - 1.062}{0.023}
\]  

(4)

This approach estimated that participants reached the PIB+ threshold an average of 8.23 (SD = 7.20, range = 0.11 to 29.5) years prior to the observed intercept PIB+ scan.
Figure 16. Plot of PIB measurements for participants with at least one PIB measurement and one UCR measurement plotted by the number of years away from PIB baseline. Points represent single PIB measures and measures collected from a single participant are connected by lines. The dotted horizontal line reflects the boundary for PIB accumulation (1.062). Panel A includes all 145 participants regardless of whether the participants ever show PIB accumulation. Panel B includes only the PIB+ scans from the PIB+ participants. The diagonal dotted lines reflect the estimated slope of PIB accumulation such that the point of intersection of the dotted lines represents the estimated number of years prior to the PIB baseline at which a given participant is predicted to have a PIB cDVR of 1.062 (i.e., become PIB+).
**Determination of Eligible UCR Measures**

The estimated event times were then used to select the UCR measures eligible for inclusion in UCR Summary measures and subsequently, the participants eligible for inclusion in the survival analysis. Although, most UCR measures were obtained prior to the PIB baseline (as shown in Figure 15), the estimated event time for the PIB+ participants typically preceded the PIB baseline measure by several years. Therefore, only 35 of the 55 PIB+ participants had two or more UCR measures collected prior to the estimated event time.

The same eligibility criteria was applied to the UCR measures for the PIB- group. However, because the event time for this group typically occurred at, or after, the PIB baseline (and most UCR measures were collected prior to PIB baseline), in practice, all UCR measures contributed by PIB- participants were considered eligible for inclusion. Therefore, a greater proportion (80 of 90) of the PIB- participants had two or more eligible UCR measures.

**Cortisol Summary Measures**

An average of 4.32 UCR measures (range 2 – 9) spanning an average period of 9.44 years (range 1.93 to 19.24) were included in the estimation of the UCR summary measures for the PIB+ group and 5.53 (range 2 – 11) UCR measures spanning 13.33 years (range 2.54 to 29.80) were included in the estimation for the PIB- group. For each summary measure, values greater than three SDs from the mean were excluded from the regression analyses, thereby excluding one participant from the UCR Closest analysis and two participants from each of the UCR Mean and UCR Slope analyses. Notably, a single
participant had outlying values for both the UCR Mean and UCR Slope (and was therefore one of the two participants removed from each of these analyses).

Figure 17 shows the temporal relationships between the UCR collections included in the UCR summary measures and the event times.
Figure 17. Timing of UCR measurement and event time (rounded to the nearest year). Points reflect UCR measurements. Black boxes represent the excluded participant. Printed values reflect the mean and SD for each UCR Summary measure calculated after excluding the relevant outlying participants. Vertical tick marks reflect the event time for all participants (i.e., estimated time of PIB+ for PIB+ participants or censored time for PIB- participants).
Cox Regression Models

PIB+ participants had average time to event of 2.71 years (SD = 2.84, range 0.06 to 13.43) while PIB- participants had an average time to event (i.e., time to censoring) of 7.66 years (SD = 3.88 range 0.01 to 14.67). Without the outlying UCR measures included in the UCR Summary measure calculations, the PIB+ group had significantly higher UCR Closest values ($t = -2.61, p = 0.01$) and marginally higher UCR mean values ($t = -1.98, p = 0.051$) (see
Figure 17). The groups did not differ on age of last eligible UCR measure for any of the models ($p > 0.84$). None of the covariates or models violated the assumption of proportional hazard on the basis of the Schoenfeld residuals, although age at last UCR visit
suggested the potential for some dependence ($p_s = 0.13, 0.09$ and $0.12$ for the UCR Mean, UCR Slope and UCR Closest models respectively).

The results of the Cox regression models are displayed in Table 6. For the unadjusted models, both the UCR Slope and UCR Closest were significant predictors of hazard ($p = .04$ and $p = .01$, respectively). For both measures, adjusting for age strengthened the influence of the UCR measures and resulted in significant likelihood ratio tests ($p = 0.03$ and $p = 0.02$ respectively) with age at last eligible UCR measure showing a marginally significant effect ($p_s = 0.09$).

The associated hazard ratios for these models suggest that for each increase of 1 in UCR Slope, the hazard of becoming PIB+ increases by 38% and for each unit in UCR Closest, the hazard increases by approximately 4%. Age, while not significant alone, appears to increase hazard by approximately 3-4% per year of age.
Table 6: Cox Regression Analysis Results

<table>
<thead>
<tr>
<th></th>
<th>Unadjusted</th>
<th></th>
<th></th>
<th></th>
<th>Adjusted</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>b</td>
<td>HR</td>
<td>95% CI</td>
<td>p</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCR Mean</td>
<td>0.031</td>
<td>1.032</td>
<td>(0.994 - 1.070)</td>
<td>0.097</td>
<td>0.028</td>
<td>1.028</td>
<td>(0.991 - 1.067)</td>
<td>0.142</td>
</tr>
<tr>
<td>UCR Slope</td>
<td>0.273</td>
<td>1.314</td>
<td>(1.013 - 1.705)</td>
<td>0.040</td>
<td>0.320</td>
<td>1.377</td>
<td>(1.051 - 1.804)</td>
<td>0.020</td>
</tr>
<tr>
<td>UCR Closest</td>
<td>0.030</td>
<td>1.031</td>
<td>(1.006 - 1.056)</td>
<td>0.014</td>
<td>0.033</td>
<td>1.034</td>
<td>(1.008 - 1.060)</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Likelihood-Ratio Test = 3.903 (2 df) = 0.142

Likelihood-Ratio Test = 6.918 (2 df) = 0.031

Likelihood-Ratio Test = 8.247 (2 df) = 0.016

Note. b reflects the regression coefficient (log hazard) for each covariate. HR: Hazard ratio. UCR Mean and UCR Slope models included N = 113 with 34 events (i.e., PIB+ cases) and UCR Closest models included N = 114 with 35 events.
Post-hoc Analyses

Considering the significant impact of UCR Slope and UCR Closest on PIB positivity hazard coupled with the marginally significant impact of age in both adjusted models, post-hoc models were conducted by adding a UCR x Age interaction term. For both models, the interaction term was not significant, suggesting that the increased hazard conferred by elevations in UCR Slope or UCR Closest measures is not moderated by the age of last eligible UCR measure (see Table 7).
Table 7: Post-hoc Cox Regression Analysis Results

<table>
<thead>
<tr>
<th>Post-hoc</th>
<th>b</th>
<th>HR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCR Slope</td>
<td>1.101</td>
<td>3.007</td>
<td>(0.606 - 14.917)</td>
<td>0.178</td>
</tr>
<tr>
<td>Age</td>
<td>0.039</td>
<td>1.040</td>
<td>(0.997 - 1.085)</td>
<td>0.070</td>
</tr>
<tr>
<td>UCR Slope * Age</td>
<td>-0.011</td>
<td>0.989</td>
<td>(0.967 - 1.012)</td>
<td>0.331</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Likelihood-Ratio Test = 7.782 (3 df) = 0.051</td>
<td></td>
</tr>
<tr>
<td>UCR Closest</td>
<td>0.015</td>
<td>1.015</td>
<td>(0.836 - 1.232)</td>
<td>0.878</td>
</tr>
<tr>
<td>Age</td>
<td>0.027</td>
<td>1.027</td>
<td>(0.948 - 1.113)</td>
<td>0.516</td>
</tr>
<tr>
<td>UCR Closest * Age</td>
<td>0.000</td>
<td>1.000</td>
<td>(0.997 - 1.003)</td>
<td>0.854</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Likelihood-Ratio Test = 8.281 (3 df) = 0.041</td>
<td></td>
</tr>
</tbody>
</table>

Note. \(b\) reflects the regression coefficient (log hazard) for each covariate. UCR Slope model included \(N = 113\) with 34 events (i.e., PIB+ cases) and UCR Closest model included \(N = 114\) with 35 events.
Discussion

The results of this analysis suggest that individuals with increasing basal cortisol levels during mid- to later-adulthood are at increased risk for the initiation of Aβ accumulation in the cerebral cortex compared to adults with stable or decreasing cortisol levels. Similarly, higher basal cortisol levels are associated with an increased risk of Aβ accumulation initiation during the following decade. In line with earlier findings, advancing age independently increased the risk of initial Aβ increase (Koscik et al., 2020; Lim et al., 2017; Vlassenko et al., 2011), however, the increased risk conferred by basal cortisol increases or elevations was not moderated by age.

Before discussing these results as they pertain to the relationship between basal cortisol levels and Aβ deposition, it is important to consider the underlying assumptions about the process of Aβ accumulation and the biological interpretation of the observed basal cortisol levels. In regards to the process of Aβ accumulation, the linear mixed effects model used here to estimate event time, as well as several studies utilizing longitudinal in vivo Aβ imaging (Burnham et al., 2020; Sojkova et al., 2011; Vlassenko et al., 2011) demonstrated that cerebral Aβ deposition is largely a unidirectional process with little evidence for a plateau until late in the progression (Jack et al., 2013). This positive, generally linear increase is indicative of a trigger scenario, whereby a single event or condition provokes excessive accumulation which continues consistently, rather than gradual or fluctuating onset of Aβ accumulation.

ALH Framework

Although the GCH discussed earlier outlines a mechanism which can account for the basal cortisol elevations commonly observed in AD, the GCH is relatively limited in
scope. That is, the GCH focuses primarily on the HPA axis and the mechanism by which the self-regulatory nature of the system can fail in pathological aging (Sapolsky et al., 1986). Given that cortisol and Aβ interact in the context of multiple biological systems, often in the absence of true hypersecretion of cortisol, and, arguably, even in the absence of pathological aging, the present findings consider basal cortisol elevations and the reported interaction with Aβ in the context a broader but related framework: the ALH (McEwen, 1998).

The ALH outlines that when presented with adversity (physical or psychological) several biological systems are forced to adapt including the autonomic nervous system, metabolic systems, immune system and the HPA axis. Short-term adaptations (such as increases in cortisol, cholesterol, or blood pressure) are necessary to address the experienced challenge and restore homeostasis, which, is ultimately beneficial. However, these adaptations put “wear and tear” on the body, and this cost is termed “allostatic load”.

Generally, these biological systems are able to efficiently manage the return to homeostasis after encountering adversity. However, dysregulation of the adaptive systems can manifest in several ways, including the failure to mount a sufficient response to or hyper-responsivity to experienced adversity, or, as is relevant for basal cortisol elevations, failure to downregulate after returning to a homeostatic state. Therefore, in the context of the ALH small chronic elevations in basal cortisol, like the chronic elevations in cholesterol or blood pressure common in aging, reflect minor biological dysregulation (McEwen, 1998), rather than the severe, pathological elevations.
The pattern of results from the survival analysis are consistent with an ALH-based interpretation of basal cortisol levels and a trigger-scenario of Aβ accumulation. In the context of the ALH, a positive UCR Slope arguably reflects a situation wherein, over a period of years, the HPA-axis increasingly fails to downregulate after the experience of adversity resulting in gradually increasing basal cortisol levels. In turn, these increasing basal cortisol levels put more “wear and tear” on the body, that is, they contribute more allostatic load. The increased risk of PIB positivity in these positive-slope individuals suggests that some threshold of allostatic load was more often reached in these individuals, subsequently inhibiting proper production or clearance of cerebral Aβ, thereby triggering accumulation.

A similar explanation applies to the significant increase in PIB positivity risk conferred by higher cortisol levels measured closest to event time. That is, the individual is experiencing elevated cortisol-based allostatic load and is therefore more likely to reach the allostatic load threshold necessary to trigger Aβ accumulation. In addition to increasing the risk of Aβ positivity indirectly by increasing allostatic load, cortisol elevations are hypothesized to contribute directly to increases in cerebral Aβ. Green et al. (2006) exposed cultured mouse neuronal cells to several levels of glucocorticoids and reported an increase in the secretion of Aβ40 and Aβ42 in a dose-dependent manner. If this finding extends to living humans, it suggests that cortisol elevations may contribute to dysregulation of Aβ both by diminishing the biological resources for regulating Aβ metabolism and clearance and by directly increasing neuronal secretion of Aβ oligomers.

Notably, the UCR Mean, reflecting the average basal cortisol level over a period of years to decades, did not significantly impact PIB positivity risk. It is speculated that
this null relationship reflects the need for a change in allostatic load to trigger excess Aβ aggregation. Therefore, an individual may demonstrate consistent basal cortisol elevations for a period of years or decades but never trigger excessive Aβ accumulation if adaptive dysregulation is only evident in the HPA-axis system. Alternatively, averaging across the length of cortisol history could introduce noise. That is, while a twenty-year cortisol history is informative, UCR measures collected two decades prior to the trigger of Aβ increases may not contribute meaningfully to the assessment of allostatic load or relative change in allostatic load nearer to the event time.

Finally, the marginally significant increase in hazard conferred by age of last eligible UCR measure is consistent with the mixed findings in the literature (Lim et al., 2017; Murphy, K. R. et al., 2013). Interpreted in the context of the ALH, age can be considered a very rough approximation for experienced adverse events. That is, older individuals are likely to have encountered more biologically challenging events over their lifetime and therefore have more “wear and tear” biologically. However, a multitude of experiences and lifestyle factors can contribute to the experienced allostatic load, suggesting that age is easily measured but extraordinarily rough approximation of allostatic load. This is further evidenced by the null UCR x Age terms in the post-hoc models. If age accurately reflected carried allostatic load, age would be expected to exacerbate any negative influence of the UCR measures.

In addition to being driven by allostatic load, it is conceivable that Aβ accumulation itself is a marker of allostatic load. A recent review highlighted the potential regulatory physiological functions of Aβ including promotion of recovery from brain injury, regulation of synaptic function in the hippocampus, and maintenance of the
blood-brain barrier (Brothers et al., 2018). The potential role of Aβ in the maintenance of the blood-brain barrier is particularly interesting given that the blood-brain barrier has been proposed as a mechanism through which the APOE e4 allele exerts a negative influence on cognition (Montagne et al., 2020). The blood-brain barrier is also credited with preventing bound circulating cortisol from accessing hippocampal receptors (Mendel, 1989).

If one interprets Aβ accumulation as a marker of allostatic load, the association with cortisol levels could be coincidental rather than causal. Specifically, if similar adverse events trigger both the HPA axis and, challenge blood-brain-barrier integrity or cause synaptic dysregulation, the experience of these events would simultaneously produce cortisol elevations and increase Aβ presence.

Given prior rodent work demonstrating that the injection of Aβ resulted in increased glucocorticoid levels (Brureau et al., 2013; Pineau et al., 2016; Zussy et al., 2013), it is even plausible that Aβ is driving the observed increases in UCR, rather than increased UCR driving Aβ accumulation risk. This is an important alternative to consider given both the crude estimation of event time and the methodological limitations of PIB measurement. Specifically, the estimated event time, in itself a theoretical approximation of the hypothesized trigger point, was calculated using a linear model which explicitly did not attempt to model individual differences in PIB trajectory. Therefore, it is likely that the estimated event times are less than perfect descriptions of when participants had Aβ accumulation equivalent to an observed cDVR of 1.062.

Further, the 1.062 threshold reflects the minimum reliably detected cerebral Aβ deposition using PIB rather than a strictly biological definition of abnormal or increased
accumulation. Therefore, while it is assumed on the basis of imaging that Aß deposition is triggered then accumulates linearly, it is possible that sub-PIB-detection increases from baseline begin months to years prior to the event time calculated here. Such a scenario would challenge the analysis here as all included UCR measures are hypothesized to reflect cortisol levels prior to the triggering event.

**Analytic Decisions and Assumptions**

**Linearity of PIB Accumulation**

A vital assumption of the present analytic approach was that increases in cerebral Aß accumulation, and therefore the observed cDVRs of PIB are linear. This assumption has two major implications for the present design. First, individuals can only transition from PIB- to PIB+, and participants can only transition between states a maximum of one time (Singer & Willett, 2003, p. 310). The examination of Figure 16 demonstrates that within the observed data this assumption was true for all cases. That is, no participant had a scan classified as PIB+ then a later scan classified as PIB-.

Second, the estimated event times for PIB+ participants were determined assuming a linear increase in the cDVR of PIB. Visual inspection of cDVR trajectories do not strongly suggest non-linearity, however, there was a small negative correlation ($r = -0.25$) between the PIB intercept and slope. That is, as the cDVR from a participant’s first PIB+ scan increased, the steepness of the slope decreased slightly. This decrease in slope could be reflective of a plateauing in Aß deposition, as some have hypothesized (Jack et al., 2013). Alternatively, it could reflect a shaper increase in the earlier stages of Aß accumulation. If the assumption of linearity is incorrect, it could produce systematic inaccuracy in the estimated event times, underestimating the event time (i.e. estimating a
younger age at PIB+ than actually occurred) for participants with lower cDVR intercepts and overestimating event times for participants with high cDVR intercepts.

A related assumption is that there are not significant between-person differences in the trajectory of PIB cDVR. This assumption is supported by the small amount of between-person variance in PIB slopes here, as well as other studies which have reported little variance in PIB Aβ accumulation trajectories (Sojkova et al., 2011). Like the assumption of linearity, the violation of this assumption would result in systematic error in the estimation of event times. Fortunately, Cox regression analysis relies on the order of event time, disregarding differences in spacing between events. Given the spread of the estimated event times, minor inaccuracies in the estimated event time are likely to only result in very minor inaccuracies in the true order of event occurrence.

Low PIB+ Threshold

While there is some disagreement in the literature about the appropriate thresholds for the designation of PIB positivity (see Villeneuve et al., 2015) a low cutoff (cDVR of 1.062) was used here due to the focus on PIB accumulation without the intention to make diagnostic predictions. The 1.062 benchmark used here is believed to reflect the minimum level of reliably detectable Aβ accumulation when using the scanning procedures described earlier (Lopresti et al., 2005). The use of this low threshold maximized the number of events (i.e., participants with at least one PIB+ scan) but does not provide good diagnostic specificity.

Interpretation of UCR Measures

Also fundamental to the interpretation of these results are the assumption that UCR measures reflect basal cortisol levels and that basal cortisol levels (and therefore
UCR measures) are reasonably stable over time within-person. Although 24-h UCR measures are not impervious to short-term stressors encountered on the day of collection, the noise introduced by moment-to-moment variations is reduced. Further, in the context of the 24-h urine samples collected from BLSA participants specifically, measures were collected in a controlled, clinical setting. This approach helps to minimize between-subject differences in the experience of stressors on the day of collection, although obviously not eliminating between-subject differences in general stress experience.

Regarding the within-person stability of 24-h measures, using a sample of over 900 participants, Rosmalen et al. (2014) showed that 24-h urinary free cortisol showed high ($r \approx 0.7$) within-person stability when collected over two consecutive days and moderate stability ($r = 0.6$) when pairs of two-consecutive-day averages were compared an average of two years apart. Notably, shorter follow-up intervals were associated with higher between-occasion correlations. Similarly, van Ockenburg et al. (2016) reported that the intra-class correlation between the average of daily UFC measures in two consecutive months was 0.93.

The stability of 24-h UCR measures and the validity of interpreting within- and between-subject differences in these measures is fundamental to this analysis due to the length of time between the last eligible UCR measure and event (or censoring) time. That is, the last eligible UCR measure (i.e., the UCR Closest measure) was collected an average of 2.71 years prior to event time. The significant increase in conferred risk as a function of the UCR Closest level is only meaningful if one assumes that this measure reflects the typically experienced cortisol level for a given participant around the time of
collection and that between-subject differences in measured UCR reflect between-subject differences in cortisol experience.

**Exclusionary Criteria**

Due to the focus of this investigation on whether cortisol levels, regardless of etiology, impact the risk of becoming PIB positive, participants were not excluded on the basis of medical conditions. Similarly, a high threshold for the exclusion of participants with outlying UCR summary measures was implemented in an effort to balance the tradeoff between examining the full spectrum of observed basal cortisol levels and insuring that the results were not driven entirely or exclusively by a handful of influential cases.

**Covariate Selection**

Age was included as a covariate, rather than the metric of event time as is often done in epidemiological research due to constraints of the Cox regression model coupled with the structure of the available BLSA data. Specifically, while the Cox regression model can include time-varying covariates, it requires that all participants have available measures for all observed event times. Therefore, in order to use age as the event time here, UCR measures and PIB status would have to be known for all participants at all observed event times. Given the time unstructured nature and varying ages of entry into the BLSA, this information was not present in the data.

In order to circumvent this methodological limitation, all participants were aligned to a single pre-event time (i.e., last UCR measure prior to the event). This forced all included UCR measures to be collected prior to the start of time, allowing them to be considered time invariant (despite the UCR Mean and UCR Slope conveying information
regarding longitudinal change in cortisol). Unfortunately, this manipulation renders the start of time conceptually meaningless.

No other covariates were examined due to the small sample size (i.e., the inclusion of only 35 events). On the basis of the ALH, as well as cellular and rodent studies, other measures of allostatic load (such as cholesterol or hypertension) are likely to be associated with Aβ accumulation (Ashby et al., 2016; Gamba et al., 2012; Gomez et al., 2018) or even moderate the observed risk conferred by cortisol elevations. However, any analysis comprehensive enough to truly examine the independent effect of cortisol, is well beyond the data capability here.

**Strengths**

The strengths of this analysis stem largely from the structure of the BLSA and the data collection protocols. Specifically, the longitudinal nature of the cortisol collection (including exceptionally long follow-up periods) and fortuitous pattern of UCR measurement and PIB scanning were imperative to this analysis. As shown in *Figure 15*, several included participants have upwards of 20 years of UCR measures and most participants having at least ten years of measures. Thus, the available length of cortisol history greatly exceeds most studies of basal cortisol. Further, had the BLSA obtained UCR measures only in tandem with PIB measures (and not for the preceding three decades), the estimated event times for the PIB+ participants would have almost all occurred well before the measurement of any UCR measures.

In addition, the collection protocol for basal cortisol levels fully leveraged the advantages available to large-scale studies. Specifically, 24-h urine collections were obtained in a clinical setting with supervision, likely improving participant compliance to
protocols. The use of 24-h UCR measures minimizes noise contributed by the influence of circadian fluctuations in cortisol secretion producing a measure of basal cortisol levels that is more reliable than cortisol measures derived from plasma or salivary samples (Rosmalen et al., 2014; van Ockenburg et al., 2016). Further, 24-h creatinine levels were simultaneously assayed allowing for the exclusion of obviously incomplete collections (Barr et al., 2005; De Keyzer et al., 2012; Kacsoh, 2000; Wu, A. H. B., 2006). Then, the ratio of free cortisol to creatinine was used to control for age-related reductions in kidney functioning (Alessio et al., 1985; Timiras, 1995). Finally, all samples were assayed for free cortisol using liquid chromatography with mass spectrometry, a method which does not suffer from the same overestimation of free cortisol due to cross-reactivity with cortisol metabolites present in urine as radioimmunoassay techniques (Woods et al., 2008).

**Future Directions**

Future studies of the influence of cortisol on the accumulation of Aβ would benefit from focusing on targeted time periods for examination. Efforts to narrow the focus to more relevant time periods could be aided by the measurement of cortisol from hair samples, which are thought to reflect integrated cortisol levels over one to three months (Liu & Doan, 2019). This approach would be particularly useful in attempts to determine if these is a gradual increase in cortisol levels over a period of years preceding Aβ accumulation. Alternatively, the use of ambulatory or at-home assessment of cortisol (Hogenelst et al., 2019) could make regular assessment of cortisol measures more plausible on a large scale.
Conclusions

Independently of age, increases in basal cortisol levels are associated with a small but significant elevation in the risk of the accumulation of cerebral Aβ. These results further suggest that cortisol may be a viable avenue for therapeutic delay of AD progression.
CHAPTER 4: GENERAL DISCUSSION

Taken together these results support the longstanding hypotheses that HPA-axis dysregulation is involved in the etiology of AD. The meta-analytic results showed that basal cortisol levels are elevated in AD compared to cognitively normal controls. Importantly, the results from the meta-analysis also showed that there is significant variability in magnitude or even presence of hypercortisolism in the AD samples across studies. This between-study variation was not accounted for by demographic or disease characteristics including age, sex, or clinical severity which have been hypothesized to be possible important moderators of the cortisol-AD association. If cortisol is a driving component of AD, the origin of this heterogeneity could provide insight into the mechanism through which cortisol contributes to AD and the Alzheimer clinical syndrome.

The survival analysis from Chapter 3 showed that cortisol elevations often precede, and increase the risk of, initial Aβ accumulation. Although Aβ accumulation is not specific to AD and the relationship between Aβ and the Alzheimer clinical syndrome is not fully elucidated, this finding still has important implications for AD and the clinical symptoms. Specifically, if one assumes the amyloid-first progression of AD, knowledge of those factors which increase the risk of Aβ accumulation provides an avenue for therapeutic treatment long before the emergence of clinical symptoms, regardless of whether these symptoms are AD specific or not.

Taken together, these results suggest that basal cortisol elevations are involved in the pathophysiology of AD. Further, these results are in line with the extension of
findings from the in vitro and rodent literature, showing that glucocorticoids directly increase Aβ (Green et al., 2006; Wang, Y. et al., 2011). The present findings demonstrated that in humans, increases in cortisol within-person over a period of years increased the risk of Aβ accumulation (as indicated by PIB positivity). Notably, once this accumulation begins, further accumulation of Aβ (without targeted intervention) appears inevitable (e.g., Burnham et al., 2020; Mishra et al., 2018).

Further extending the in vitro and rodent work, the results of the meta-analysis are consistent with the hypothesis that Aβ toxicity (Reiss et al., 2018) in the hippocampus is exacerbated by the presence of glucocorticoids (Goodman et al., 1996). This is evidenced by the accelerated rate of hippocampal structural decline reported in AD patients (Jack et al., 1998; Zhao, W. et al., 2019) as well as the findings of the current meta-analysis demonstrating that basal cortisol elevations are present in samples with symptoms of the Alzheimer clinical syndrome. Taken together, these results suggest that therapeutic interventions targeting cortisol or relevant receptors (Canet et al., 2019a; Canet et al., 2020) have the potential to delay or prevent AD progression, particularly if interventions are implemented prior to initial Aβ accumulation.

**Alternative Interpretations**

**Directionality of Effects**

The analytic approaches and interpretations in this paper have relied on an amyloid-first model of AD. That is, it was assumed that Aβ accumulation is the initial event in the development of AD and that cognitive impairments are a downstream consequence of this accumulation (Hardy & Higgins, 1992; Hardy & Selkoe, 2002; Selkoe & Hardy, 2016). However, three recent investigations using ADNI participants
have suggested that subtle cognitive difficulties precede (Hadjichrysanthou et al., 2020) and help to predict future Aβ accumulation (Elman et al., 2020; Thomas, K. R. et al., 2020). Specifically, these researchers noted that participants showing subtle impairments on composite measures of verbal episodic memory were more likely to transition from an Aβ- to an Aβ+ state during an average follow-up of 2.8 years (Elman et al., 2020) and showed accelerated Aβ accumulation over a period of 48 months (Thomas, K. R. et al., 2020) compared to cognitively normal participants.

These findings are particularly interesting considered in the context of basal cortisol elevations and the role of the hippocampus in both episodic memory and HPA-axis regulation. Declarative episodic memory performance in particular has been the focus of cortisol research due to age-related declines in this domain (Harada et al., 2013) and limited research suggesting that basal cortisol elevations may be present in normal aging (Gupta & Morley, 2014; Purnell et al., 2004; Zaidi et al., 2012).

A handful of reports of a negative relationship between basal cortisol levels and episodic memory performance have emerged from this literature. Specifically, Comijs et al. (2010) reported a negative relationship between morning cortisol levels and immediate verbal recall, Gerritsen et al. (2011) reported a negative relationship between bedtime salivary cortisol measures and delayed verbal memory performance in a sample of 911 older adults, and Lee et al. (2007) reported a negative relationship between the AUC of four saliva measures and a composite score of immediate and delayed verbal memory performance.

While speculative, these results suggest that basal cortisol elevations could contribute to the appearance of subtle cognitive impairment and concurrently serve to
trigger Aβ accumulation. Alternatively, a third variable (or set of variables) could produce both subtle cognitive impairments and Aβ accumulation, in turn elevating cortisol levels in the accompanying time frame (Brureau et al., 2013; Pineau et al., 2016; Zussy et al., 2013).

**Uniformity of effects**

In addition to interpreting the cortisol effects in the context of the amyloid-first model of AD, the present findings, particularly the meta-analytic results, were interpreted as evidence of consistent, small to moderate cortisol elevations among most or all of the AD sample as opposed to extreme elevations of just a few AD patients. However, there is some suggestion of the presence of extreme cortisol elevations only in a subset of AD patients. For example, Swanwick et al. (1998) reported that six of 30 AD patients showed basal cortisol levels greater than two SDs from the mean of the NCC group, and that the AD sample as a whole had higher basal cortisol levels than the NCC group. However, it was not clarified whether this significant difference was maintained if the six extreme patients were removed from analyses.

The residual heterogeneity present in the moderated analyses further suggests that larger elevations limited to a subset of AD participants may be a concern. That is, if effects are in fact being driven by a small number of AD patients with extreme cortisol elevations, one would expect to see heterogeneity in the observed effect size on the basis of whether or not a sufficient number of these patients were included (and/or if the primary study used a log transformation or excluded outliers, information which was rarely reported).
However, examination of the BLSA-provided effects included in the meta-analysis do not support this concern. While the UCR distributions for both the AD and NCC samples are positively skewed, there is no evidence for extreme outliers or a bimodal distribution among the AD samples. Similarly, to my knowledge, no reports of a bimodal distribution of basal cortisol levels among AD patients exist despite the large body of research examining disease-related patterns in cortisol levels.

**Etiology of Cortisol Elevations**

Neither the meta-analysis nor survival analysis included here attempted to identify the source of cortisol elevations despite the reasonable argument that a third variable could contribute to cortisol elevations, Aβ accumulation and/or AD risk. This intentionally limited scope was taken due to the extensive body of work identifying a multitude of influences on both cortisol levels and AD risk.

For example, the experience of chronic stress is a proposed risk factor for the development of AD (Canet et al., 2019b; Caruso et al., 2018; Dong & Csernansky, 2019; Escher et al., 2019; Joshi & Praticó, 2013). While cortisol elevations are a well-established aspect of the stress response, a plethora of individual differences, ranging from pre-natal hormone exposure to late-life medication usage, contribute to the amplitude and duration of the evoked cortisol secretion as well as the neurophysiological vulnerability to these elevations. (Lucassen et al., 2013; Lupien et al., 2018; Lupien et al., 2016; Wippert et al., 2014).

Not only did the present analyses lack the statistical power to systematically evaluate all measured potential covariates, the potential impact of dozens of other unmeasured covariates could be speculated about. However, given the mixed reports of
basal cortisol elevations in the Alzheimer clinical syndrome and the lack of research on the temporal relationship between cortisol elevations and initial Aβ accumulation, the goal of these investigations was to determine if cortisol elevations are characteristic of AD or predictive of PIB positivity, not to investigate whether elevations exert independent effects on these factors.

Methodological Considerations

Unidimensionality of Basal Cortisol

Throughout this paper, basal cortisol levels have been referred to as a unitary entity and little distinction has been made between findings on the basis of cortisol sampling methodology. Although all of the sampling and assay protocols used in the studies of the cortisol- Aβ and cortisol-AD interactions considered here are valid, a number of methodological considerations suggest that caution should be taken when interpreting cortisol results across studies.

In the context of the relationship with AD and Aβ, the goal of cortisol measurement in other biological fluids is often to estimate neural exposure. Qian et al. (2012) reported significant cross-correlations between free glucocorticoids sampled from the blood and the hippocampus simultaneously in rodents taken at 10 or 30 minute increments for 20 hours. This finding suggests that neural exposure to glucocorticoids follows the same circadian rhythm that would be expected on the basis of free hormone blood sampling.

However, there is limited and mixed support regarding whether this alignment extends to humans. Specifically, both Popp et al. (2015) and Laske et al. (2009) measured cortisol levels from consecutively collected blood and CSF samples. Popp et al. (2015)
reported a significant correlation ($r = 0.31$, $p < 0.001$) between the two levels while Laske et al. (2009) did not find a significant association ($r = 0.14$, $p = 0.50$).

Further complicating the interpretation of cortisol levels is the low day-to-day reliability of cortisol sampled using single time-point collections (such as blood or saliva). Zhang et al. (2017) reported that salivary samples collected one week apart at waking, noon, and 17:00 were not significantly correlated with each other (ranging from $r = 0.11$ to $r = 0.18$). Urinary cortisol measures have been shown to be more stable with significant correlations between measures collected a week apart (correlations between $r = 0.57$ and $r = 0.67$ reported by Short et al. (2016) and $r = 0.29$ and $r = 0.32$ reported by Zhang et al. (2017)). The low correlations between single time-point measures suggest that these measures may not provide a consistent within-person estimation of basal cortisol levels.

**PIB and Revisions to the Amyloid Cascade Hypothesis**

Despite remaining at the forefront of AD research, the initial Amyloid Cascade Hypothesis (Hardy & Higgins, 1992) has been widely criticized as evidence accumulates that direct relationships between Aß plaque density and dementia severity are often not present (Arriagada et al., 1992; Gómez-Isla et al., 1997; Nagy et al., 1995). A number of modifications and extensions to the original hypothesis have been made to explain this discrepancy (Hardy & Selkoe, 2002), including the Oligomer Hypothesis (Lambert et al., 1998), and a ratio hypothesis (Jan et al., 2008; Kuperstein et al., 2010; Walsh & Selkoe, 2007). The Oligomer Hypothesis proposes that it is only the soluble oligomers (i.e. the polymers of aggregated Aß that consists of only a few repeating units) rather than Aß plaques which contribute to neural degeneration. Similarly, the ratio hypothesis posits...
and increase in the ratio of Aβ_{42} to Aβ_{40} is the defining and damaging trait of AD rather than plaque accumulation alone.

These revisions are of practical importance for the interpretation of PIB results. PIB, like thioflavin-T, has a high affinity for *aggregated* Aβ. Therefore, it does not provide any estimate of the presence or density of Aβ oligomers (Perani et al., 2019), nor does it provide any indication of the ratio of Aβ_{40} to Aβ_{42}. Thus, while PIB is a valid measure of fibrillar Aβ accumulation it may not be a linear marker of AD severity.

Fortunately, reports from in-vivo imaging analysis do suggest that the cortical pattern and staging of amyloid deposition measured in vivo reflect those previously observed post mortem (Grothe et al., 2017; Thal et al., 2002) and this staging may provide information about future clinical changes.

**Cautionary Notes**

While these results support the involvement of basal cortisol elevations in AD it is crucial to note that there is no evidence to suggest that basal cortisol elevations are the sole cause of Aβ accumulation or the cognitive deficits associated with the Alzheimer clinical syndrome. First, the average effect size as calculated by the meta-analytic approach is fairly small in contrast to other conditions with basal cortisol elevations. For example, Viardot et al. (2005) reported that Cushing’s disease patients had 24-h UCR measures 16 times higher than normal controls. Meta-analyses of depressed patients in comparison to normal controls have reported effect sizes more than double those observed here, however, the effect sizes are moderated by methodological differences (Murri et al., 2014; Stetler & Miller, 2011).
Notably, although patterns of hippocampal structural damage, as well as impairment and accelerated decline on hippocampal-dependent tasks mirroring those of AD are observed in both Cushing’s syndrome (Dorn & Cerrone, 2000; Forget et al., 2016; Forget et al., 2000; Martignoni et al., 1992; Mauri et al., 1993; Starkman et al., 2001; Whelan et al., 1980) and depressed patients (or those with depressive symptoms) (Barnes et al., 2006; Geerlings & Gerritsen, 2017; Rubinow et al., 1984; Scult et al., 2017; Snyder, 2013), the severity of these symptoms is milder than typically observed in AD or the Alzheimer clinical syndrome. This suggests that while basal cortisol elevations may exacerbate structural and functional declines, some other mechanism or mechanisms characteristic of AD are responsible for the severity and progression of AD.

Similarly, Aβ accumulation is not specific to AD (Catafau & Bullich, 2015) and is not uncommon even among elderly showing no evidence of significant cognitive impairment (Jansen et al., 2015). Therefore, while there is little dispute that Aβ accumulation is characteristic of AD or that Aβ negatively impacts cognition (Baker et al., 2017), it is not an accurate to surmise that the initiation of Aβ accumulation is evidence of imminent transition to AD.

Conclusions

The combined findings reported here demonstrate that basal cortisol elevations are a frequent feature of the Alzheimer clinical syndrome and cortisol elevations increase the risk of Aβ accumulation. These results suggest that therapeutic interventions targeting cortisol may be able to delay the onset of the disease and corresponding clinical syndrome.
REFERENCES


Aoki, S. (2020). Effect sizes of the differences between means without assuming the variance equality and between a mean and a constant. *Heliyon, 6*(1), e03306. doi:10.1016/j.heliyon.2020.e03306


119


121
circulating and genetic factors in association to Alzheimer’s type dementia. *Clinical biochemistry, 42*(9), 783-790. doi:10.1016/j.clinbiochem.2009.02.006


149


