Olfactory Identification Decline: A Preclinical Biomarker for Alzheimer's Disease

by

Jamie Knight
B.A., University of Victoria, 2015

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Psychology

© Jamie Knight, 2017
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

Supervisory Committee

Olfactory Identification Decline: A Preclinical Biomarker for Alzheimer's Disease

by

Jamie Knight
B.A., University of Victoria, 2015

Supervisory Committee

Dr. Andrea Piccinin (Department of Psychology)
Supervisor

Dr. Mauricio Garcia-Barrera (Department of Psychology)
Departmental Member

Abstract

Supervisory Committee
Dr. Andrea Piccinin (Department of Psychology)
Supervisor

Dr. Mauricio Garcia-Barrera (Department of Psychology)
Departmental Member

The earliest stage of Alzheimer's disease (AD) pathology begins in one of the main components of the olfactory pathway, the entorhinal cortex, making deficits in smell a potential prospective biomarker for the early detection of AD. This study contributes to the field with a reproduction and extension of work by Wilson, Arnold, Schneider, Boyle, Buchman, and Bennett (2009). A sample of 1089 individuals (M=78.4 years), more than double the data available in 2009, completed annual assessments of olfactory functioning, and cognitive functioning for up to 18 years with brain donation at death. Mixed effects models conditioned on demographics estimated between and within-person effects in olfactory functioning and episodic memory (EM). After successful reproduction of Wilson et al. (2009), addition of AD pathology (ADP) demonstrated that both ADP and olfaction were significantly related to EM at baseline. Higher ADP at autopsy was significantly related to faster declines in olfaction, as well as more rapid declines in EM. Higher olfactory scores were associated with higher EM scores and a model for EM with olfaction as time-varying covariate indicated that at a given occasion, individuals with higher olfactory scores also have higher EM scores.

These results align with the hypotheses that difficulty in identifying odors predicts development of cognitive impairment; increased levels of AD pathology are related to both decreased EM at baseline and faster declines, as well as faster rates of decline in olfaction; and olfaction and cognition are travelling together over time.

Keywords: dementia, neurodegenerative, disease, olfaction, smell, cognitive decline, Alzheimer's

Table of Contents

List of Tables

List of Figures

Acknowledgments

This work would not have been possible without the support of many people.

First and foremost, I would like to thank my family. Their love and support gave me the motivation to stay engaged through the most arduous parts of this journey. The journey, in the most literal sense, would not have been possible if not for Al and Nancy who took me in to ease the burden of the long commute. Thank you for taking such good care of me over the past 2 years.

Without the encouragement from Dr. Elliott Marchant I would have never thought to apply to graduate school. A few well-placed words managed to completely change the direction of my life, and I am very grateful.

Deepest gratitude to my supervisor, Dr. Andrea Piccinin, who offered invaluable support and guidance. Her mentorship has ignited a keen interest in methods and statistical modeling that has channeled my passion for research.

I would also like to thank past and present members of the iLifespan Lab—Raquel Graham, Andrey Koval, Jonathan Rush, Tomiko Yoneda, and Rebecca Vendittelli—for their friendship, advice, and support. I would also like to acknowledge the University of Victoria for their graduate fellowships in support of this research.

None of this research could have been possible without the exemplary and diligent work of the teams responsible for collecting the longitudinal data for the Rush University Memory and Aging Project, and the participants who generously donated their time and brains to science.

Thank you.

Dedication

For Greyson, Vesper, and James:

“Two roads diverged in a wood, and I—

I took the one less traveled by,

And that has made all the difference.”

- Robert Frost

Introduction

Smell plays an important role in health and behaviour. It is critical for maintaining appetite, identifying food, avoiding hazards and pathogens, finding a mate, and sustaining intimate social relationships (Croy, Nordin, & Hummel, 2014; Murphy 2008). Maintaining a normal sense of smell requires the constant cellular regeneration of the olfactory epithelium, olfactory bulb, and hippocampus (Lledo, Alonso, & Grubb, 2006; Whitman & Greer, 2009). The loss or malfunction of these regenerative processes will cause deficits in the ability to detect or identify smells (Doty, 2009; Shepherd, 2007). The earliest stages of Alzheimer's disease pathology begin in one of the main components of the olfactory pathway, the entorhinal cortex, indicating deficits in smell as a prospective biomarker for the early detection of Alzheimer's disease (AD; Albers, Tabert, & Devanand, 2006).

The purpose of this project is three-fold: (1) provide a first reproduction of the design, measures and statistical approach of Wilson, Arnold, Schneider, Boyle, Buchman and Bennett (2009) and contribute to the small body of literature on longitudinal analysis of olfactory decline; (2) extend the existing model by (a) including autopsy diagnosed AD pathology; and (b) conducting a separate linear mixed model with BSIT as outcome to examine the relationship between autopsy diagnosed AD pathology and Brief Smell Identification Test (BSIT) score. The third and final aim (3) is to examine change in memory performance not only as a function of the passage of time but also as a function of (intraindividual) changes in smell: a growth model for episodic memory with BSIT as time-varying covariate (also called a coupling model). This approach assumes that smell could be a proxy for neurological damage.

The majority of work in this area has been cross-sectional. However, the few longitudinal studies in the literature are very promising (Devanand et al., 2015; Wilson et al., 2009; Wilson, Arnold, Tang, & Bennett, 2006). A *replication* is usually defined as the replication of the same methods in a different data set, whereas a *reproduction* is replicating the same methods in the same data set. The first aim of this study combines replication and reproduction, as the methods are reproduced in data from the same longitudinal study, yet the sample is much larger and cannot be considered identical to that of the original paper. With this caveat in mind, it will be referred to here as a reproduction. To the best of my knowledge, this is the first independent reproduction of this work.

Alzheimer's Disease & Pathology

Dementia is a general term for a decline in mental ability that is severe enough to interfere with daily life. Alzheimer's disease (AD), the most common form of dementia, accounts for nearly 70% of all diagnosed cases. There are two major forms of AD: (a) early onset, a genetic version, which comprises only a small percentage of diagnosed cases; (b) late onset Alzheimer's disease, which comprises the majority of cases and is presumed to have both genetic and environmental factors; the latter is the Alzheimer's that will be discussed in this paper. It is still unclear what causes late onset Alzheimer's disease (Fiandaca, Mapstone, Cheema, & Federoff, 2014) and definitive classification currently requires a brain autopsy (Wilson et al., 2009).

Symptoms of Alzheimer's disease are debilitating and include disturbed memory, difficulty sleeping, drastic mood changes and general confusion. Many of the symptoms are a result of mis-folded Amyloid-β (Aβ) oligomers (neuritic plaques) and malfunctioning tau protein polymers (neurofibrillary tangles) causing cell death and degradation of neuronal connections in the brain (Walker, Diamond, Duff, & Hyman, 2013).

Tau.

Tau is abundant in neurons and its function is to stabilize microtubules. Tau controls microtubule stability using isoforms (different variants of the tau protein) or phosphorylation (the addition of a phosphoryl group to the tau protein). Phosphorylation acts as an on/off switch, with phosphorylated tau disrupting microtubule organization. In AD pathology hyperphosphorylated tau (*p*-tau) proteins combine to form a filamented polymer, a neurofibrillary tangle, these accumulate inside neurons, killing the cell, and causing neurodegeneration (Walker et al., 2013). The tangles are thought to progress in a specific pattern triggered by cell-to-cell transmission. This was described by Braak in 1991, and is now called the "Braak stages" (Braak & Braak, 1991). Stages I & II are associated with preclinical Alzheimer's disease and are defined by the presence of neurofibrillary tangles in the transentorhinal region of the brain (Braak & Braak, 1991; Goedert, Clavaguera, & Tolnay, 2010). Stages III and IV are thought to correlate to MCI (mild cognitive impairment) and are defined by the progression of tangles through to limbic regions such as the hippocampus. Stages V and VI are associated with AD, and at this point extensive tangles have spread into the neocortex. Tau pathology can be seen at a much earlier age than neuritic plaques. Braak and Del Tredici (2011), using autopsy studies, found evidence of tau pre-tangles (i.e., abnormally phosphorylated tau proteins in the absence of mis-folded Aβ) in subcortical and brainstem nuclei and the entorhinal cortex in children as young as six (Braak & Del Tredici, 2011). Braak et al. proposed that tau pathology begins in the locus coeruleus (located in the brainstem near the 4th ventricle in the rostral pons), and it then spreads by cell-to-cell transmission through the brain (Braak & Del Tredici, 2011, 2015; de Calignon et al., 2012). This hypothesis has been criticized due to the fact that Aβ over-production leads to AD whereas a high level of initial tau does not lead to inevitable cognitive impairment nor does it lead to inevitable AD (Jack Jr et al., 2013). According to the Braak et al.'s hypothesis, olfaction would

be involved very early on because the first cells in the brain to develop neurofibrillary lesions, located in the locus coeruleus in the brain stem, have norepinephrine neurons that send projections to neurons in the main and accessory olfactory bulbs (Winberg & Porter, 1998). The Braak stages list the transentorhinal cortex as also being involved very early on in the disease pathology and this area is one of the primary components of the olfactory cortex (Purves et al., 2001).

Amyloid-*β.*

Amyloid-β (Aβ) peptide is a product of the cleavage of the Aβ precursor protein (APP) and is the primary component of neuritic plaques (O'Brien & Wong, 2011). Accumulating Aβ peptides aggregate to form oligomers which can mis-fold, and once they are in this pathogenic, mis-folded state, they can induce a chain reaction that propagates the mis-folding in the surrounding Aβ molecules, creating a neuritic plaque (Goedert, Clavaguera, & Tolnay, 2010; Walker et al., 2013). These neuritic plaques accumulate in the brain, block neuronal synapses, and contribute to the brain pathology that leads to Alzheimer's disease (Attems, Walker, & Jellinger, 2015; Liu, Kanekiyo, Xu, & Bu, 2013; Walker et al., 2013).

Apolipoprotein E.

The deposition of Aβ in the brain is related to the apolipoprotein E (APOE) gene. There are three allelic variants of the APOE gene and each produces a protein differing by two amino acids: ApoE4, ApoE3 and ApoE2 occurring in the population (in AD) in 14% (37%), 79% (59%), and 7% (4%), respectively (AlzForum, 2017; Farrer et al., 1997). The APOE ε4 allele is strongly associated with increased Aβ deposition in the brain (O'Brien & Wong, 2011) and is known to increase the risk of Alzheimer's disease, as well as other neurodegenerative diseases (Farrer et al., 1997; Liu et al., 2013). The ε4 allele occurs in approximately 40% of patients

diagnosed with Alzheimer's disease, and when heterozygous (only one in the pair), can increase the risk of Alzheimer's disease by up to four times (up to 15 times for those who are homozygous) when compared to individuals with homozygous ε3 alleles (Liu et al., 2013; O'Brien & Wong, 2011). On the other hand, the ε2 variant, which is very rare, has been found to protect against the development of late onset Alzheimer's disease (Liu et al., 2013). The ApoE lipoprotein functions as a cholesterol transporter in the brain, and current theory proposes that the effect of ApoE4 on AD risk may be largely due to its ability to bind Aβ together resulting in hallmark AD plaques (Walker et al., 2013; O'Brien & Wong, 2011). ApoE4 has also been found to contribute to neurofibrillary tangles by stimulating tau to abnormally phosphorylate (Murphy, Solomon, Haase, Wang, & Morgan, 2009; Kim et al., 2009).

There is some debate as to whether *p*-tau or Aβ comes first in the timeline of AD pathology (Braak & Del Tredici, 2015; Jack et al., 2010; Jack Jr et al., 2013). In contrast to what Braak et al. speculate about *p*-tau leading the pathology, Jack et al (2010) posit that Aβ is the trigger for the hyperphosphorylation of tau (Jack Jr et al., 2013). The Jack et al. (2010, 2013) model lists the posterior cingulate (using fluorodeoxyglucose (FDG)-positron emission tomography (PET) imaging) and the medial temporal lobe (using structural magnetic resonance (MRI) imaging) as the first brain areas affected in the Alzheimer's disease stages (appearance differing via imaging technique). The entorhinal cortex is in the medial temporal lobe and although there is disagreement between Braak et al. and Jack et al. about the temporal appearance of *p*-tau vs Aβ, both models indicate that the initial brain areas affected by AD pathology are also main areas of the olfactory pathway, and this may be why poor olfactory performance can be seen in conjunction with several neurodegenerative diseases (Mesholam, Moberg, Mahr, & Doty, 1998).

Current Outlook

The Alzheimer's Society has stated that, globally, 44 million people currently suffer some form of dementia. The cost of caring for those afflicted is estimated at 600 billion dollars. In 2050, it is estimated that the number of people affected by dementia worldwide will increase to 135 million. The challenge for researchers now is to find a way to easily screen the population and find individuals who are at high risk of developing dementia. When these high-risk individuals can be recruited into clinical trials, progress on finding treatments, and possibly a cure, may be within reach. Take the heart as an example: heart disease is much harder to treat after cardiac damage has occurred, as the damage impedes the success of treatments. However, when the precursor to cardiac disease, hypertension, is treated, treatments have much better results, often significantly reducing the chances of heart disease.

Brain damage and cell death from AD pathology are occurring, and accumulating, for ten to twenty years before any of the known symptoms of dementia appear (Elias, Beiser, Wolf, Au, White, & D'Agostino, 2000; Fiandaca et al., 2014; Mapstone, Dickerson, & Duffy, 2008; Masurkar & Devanand, 2014). This pre-clinical stage comprises almost 50% of the entire disease timeline and is where the least amount of information is being gathered. In order to discover effective treatments, and possibly even a cure, a method to detect the neurological precursors to dementia as early as possible is needed. Although detectable deficits may precede diagnoses by more than a decade, there is less consensus regarding the onset of accelerating memory decline (Sliwinski, 2003) and olfactory decline may provide a clearer picture of onset.

Background

Olfactory Testing

There are currently three testable facets of olfaction: (1) identification, correct selection of the name of the smell based on multiple choice; (2) discrimination, detection of differences between smells; and (3) threshold, detection of the presence of a smell. Olfactory deficits are present in patients with MCI, increase in severity with the progression of AD (Djordjevic et al., 2008), and indicate conversion from MCI to AD (Devanand et al., 2000).

There are two main types of olfactory tests: (1) Odour identification, scratch-and-sniff based tests, these include the University of Pennsylvania Smell Identification Test (UPSIT) as well as the tests based on it (e.g., Brief Smell Identification Test (BSIT), and Scandinavian Odour Identification Test (SOIT)); (2) odour pens (e.g., Sniffin' Sticks) which are more complicated to administer but have the ability to test all three domains of smell. Some studies indicate that using both threshold and identification in conjunction have the best sensitivity for detecting the early stages of disease (Djordjevic et al., 2008). Studies have reported that the UPSIT, BSIT, and Sniffin' Sitcks olfactory tests have good internal validity and reliability (Doty et al., 1984; Menon et al., 2013). However, the majority of the tests currently available exhibit ceiling effects and additional tests that can differentiate between high and low performing individuals with normative olfaction would improve the ability to detect olfactory differences earlier.

Literature Review: Cross-Sectional Studies

Olfactory ability differs widely from person to person due to genetic variations, sex, and experiences with environmental factors. The most common causes of olfactory impairment include head trauma, inflammation, upper respiratory tract infection, aging, and neurodegenerative disease (Doty & Kamath, 2014; Murphy et al., 2002). Less common factors

include exposure to toxic chemicals, congenital anomaly, craniofacial surgery, seizure disorders, cerebrovascular trauma, olfactory or temporal tumours, and endocrine disorders (Doty & Kamath, 2014; Murphy et al., 2002). These individual differences make cross-sectional studies using olfaction less than ideal, but do allow for exploration of valuable ideas and population prevalence.

Prevalence of Olfactory Disorders: Population.

Brämerson et al. (2004) looked at the prevalence of olfactory dysfunction in the Skövde Swedish population using the Scandinavian Odour Identification Test (SOIT; Brämerson, Johansson, Ek, Nordin, & Bende, 2004). The sample consisted of 1387 participants, aged 20 and up, and had found a 13.3% (95% CI, 11.6%-15.2%) prevalence of hyposmia, difficulty detecting odours, and a 5.8% (4.7%-7.1%) prevalence of anosmia, no sense of smell at all (Brämerson et al., 2004), combined a 19.1% prevalence of olfactory disorder. The definition of these groups was a score of 10-12 of 16 for hyposmia and a score of 9 or less for anosmia. Earlier, in 2002, Murphy et al. found that the prevalence of olfactory disorders among older adults (53-97) was 24.5%. When directly comparing the same age group in this study, Brämerson et al. (2004) found a prevalence of 32.9% (Brämerson, Johansson, Ek, Nordin, & Bende, 2004), these studies are likely both under-representing the prevalence as the SOIT in the Brämerson study and SDOIT in Murphy are only 12 and 8 item tests, respectively. Brämerson et al. (2004) also found that the prevalence of olfactory dysfunction increased with age and comorbidity with either diabetes mellitus or nasal polyps (Brämerson et al., 2004). Around 50 years of age there is a decline in olfactory scores and this decline accelerates after 70 years (Brämerson et al., 2004). Using a two-way ANOVA Brämerson et al. (2004) found that women performed better on the SOIT than men (no gender-by-age interaction).

The Brämerson et al. (2004) paper did not mention whether there was an option for indicating "I don't know" on the SOIT. In several of the other types of olfactory tests, such as BSIT, options for "I can't smell it" or "I don't know" are not offered, and this is a big limitation for test reliability. For example, a participant may guess the correct answer on an initial test, then on a follow up test they may guess incorrectly, thus providing a false impression of decline. This small detail may significantly affect reliability of testing and should be explored further. The attrition rate for the study was 27% with the majority being in the lowest and highest age categories. Considering this, along with the slightly higher rates of olfactory dysfunction found by Brämerson et al. (2004) versus Murphy et al. (2002), there could be some question as to whether the drop outs influenced the external validity regarding the prevalence and risk factors of olfactory dysfunction in the general population. Brämerson et al. (2004) looked at all residents in the population regardless of disease or health status, giving an idea of all-cause olfactory performance in a community-dwelling population. However, this leaves the question: does a decline in smell mean anything for the health of the population?

Olfaction and Mortality.

Using the National Social Life, Health and Aging Project community-dwelling data Pinto et al. (2014) looked at whether olfactory dysfunction could predict all-cause mortality in older adults. They used a non-standardized olfactory test, selecting five felt-tipped pen odourants (rose, leather, orange, fish and peppermint) in a forced choice protocol. They categorized anosmia as 4-5 errors, hyposmia as 2-3 errors and normosmia as 0-1 error. Pinto et al (2014) found that age increased the likelihood of death (OR, 1.07 [95%CI 1.05, 1.09]). However, after controlling for age, gender, education and race, anosmic individuals had over three times the odds of death when compared to individuals with normal olfactory test results (OR, 3.37, 95% CI 2.04, 5.57)

translating to 39% of older adults with anosmia, 19% with hyposmia and 10% with normal olfaction having died at follow-up (Pinto, Wroblewski, Kern, Schumm, & McClintock, 2014). In contrast, older age alone increased the likelihood of death by 7% (OR, 1.07 [95%CI 1.05, 1.09]). This study found that anosmia was a strong independent risk factor for mortality. However, results should be repeated to account for limitations. For example, using a standardized test would increase reliability and would allow ensuing results to be more comparable across other studies. Additionally, expanding the study to look longitudinally would give a better idea of the lifespan trajectory in relation to olfactory health rather than a single measurement. Further, looking at causes of death would provide an indication of which diseases are affecting those who show a decline in their smell. This study provides good cause to look further at the predictive value of olfaction and to examine underlying mechanisms as to why it may be a good predictor of mortality, or health status in general.

Olfaction and Aging.

With age comes a decrease in overall cellular regeneration (Campisi & di Fagagna, 2007) as well as an increase in the number of years of exposure to cumulative damage to the olfactory receptors such as sickness and air pollution (Doty & Kamath, 2014). Attems, Walker and Jellinger (2015) reviewed olfactory performance and aging and found previous research agrees that as individuals age the number of olfactory receptors decreases, the olfactory epithelium thins, the size of the olfactory bulb declines, and the foramina of the cribriform plate ossifies (Attems et al., 2015; Masurkar & Devanand, 2014). Attems et al. (2015) found that more than 50% of individuals between the ages of 65 and 80 years had decreased olfactory function and that number increased to 62-80% for those over 80 (Attems et al., 2015). Previously, Murphy et al. (2002) conducted a study using San Diego Odor Identification Test (SDOIT) data from the 5-

year follow-up examination (1998-2000) of the population based Epidemiology of Hearing Loss Study (EHLS) which consisted of 2491 Wisconsin residents aged 53 to 97 years. Murphy et al. (2002) found that the mean prevalence of impaired olfaction was 24.5% and impairment increased with age. In their sample, 62.5% of 80 to 97 year olds had olfactory impairment (95% confidence interval, 57.4%-67.7%) and olfactory impairment was more common among men (adjusted prevalence ratio, 1.92; 95% CI, 1.65-2.19), almost an exact match to Attems, Walker and Jellinger's later numbers. Schubert, Cruickshanks, Fischer, Huang, Klein, Klein, Pankow & Nondahl (2011) found a 3.8% overall prevalence of olfactory impairment that increased with age ranging from 0.6% in participants 35 years or less to 13.9% among those over the age of 65. Murphy et al (2002) included all individuals, even those with reported dementia, so different trajectories for declines in smell within a healthy aging population versus one with dementia are not apparent in the research. Ideally, differences in the rate of change between healthy aging and disease pathology would be explored to see if there is a sharp decline at some point or whether there is a linear decline in healthy aging. Further research examining the shape of disease trajectories using olfactory testing is warranted (Jack et al., 2012).

In contrast to the majority of studies, Mackay-Sim et al. (2006) found that, after excluding all individuals who were on medications, smoked, or had a history of nasal illness, the direct effect of "age" on olfactory function was relatively small (Mackay-Sim, Johnston, Owen, & Burne, 2006). The average decline in olfactory test scores from age 50 to 79 was smaller in the healthy group (2.1 points in women and 2.0 in men), when compared with the medicated group (3.2 in women and 5.2 for men; Mackay-Sim et al., 2006). This study used the Sniffin' Sticks olfactory test (Hummel, Sekinger, Wolf, Pauli, & Kobal, 1997), which assesses odour identification, discrimination, and threshold separately as well as giving a composite (TDI) score. It is hard to

directly compare scores in studies using different types of olfactory tests, but medication could account for much of the decline seen in other studies. A large effect here would not be surprising: 40% of all prescription medications reach cells via G protein-coupled receptors, also known as olfactory receptors (Filmore, 2004). The Mackay-Sim et al.'s study (2006) was cross sectional and these findings should ideally be replicated in a longitudinal study.

Cross-sectional methods provide insight into the population prevalence and the overall trend for smell to decline with age. However, longitudinal methods are required to establish intraindividual timelines of what olfaction and disease trajectories look like in relation to one another and will provide a clearer picture of which patterns indicate disease.

Prevalence of Olfactory Disorders: Disease.

Prevalence of olfactory dysfunction has been reported as high as 100% in Alzheimer's disease (Duff, McCaffrey, & Solomon, 2002), 90% in Parkinson's disease (PD); 96% in Frontotemporal dementia (FTD) and 15% in Vascular dementia (VaD; Alves, Petrosyan, & Magalhães, 2014).

Duff et al. (2002) examined the specificity and reliability of olfactory testing for differentiating between AD (n=20), depression (n=20) and Vascular dementia (n=20; Duff, McCaffrey, & Solomon, 2002). Using the Pocket Smell Test, a three-item olfactory test based on the UPSIT, they found 100% sensitivity, 92.5% specificity, and 95% overall classification accuracy (Duff et al., 2002). There were no false positives or negatives in the AD group and 3 false positives in the VaD group (3 VaD classified as AD). In the Duff et al. (2002) study there were a few limitations, the largest of which was that the diagnoses of AD were not backed up by PET scan or autopsy and potential misdiagnoses in the sample could affect the results. Other limitations include the small sample size and the very limited 3-item forced-choice olfactory test.

The pathological features for vascular dementia, depression and AD are different, and being able to differentiate between them with olfaction is a sensible hypothesis; what is unexpected is that Duff et al. (2002), using only a three item olfactory test, were able to identify AD with such high accuracy. Replication of this study using autopsy diagnosis of AD and a more in-depth olfactory test would further support these conclusions. Vascular dementia and AD have different disease pathologies, and a follow-up study comparing Parkinson's, AD, and FTD using an olfactory test with more items and an "I don't know" option, would give insight into whether olfactory testing can differentiate between diseases with similar brain pathologies. Establishing whether olfactory trajectories differ by disease is an alternative approach with potential diagnostic value.

Limitations.

The variety of potential causes of olfactory dysfunction has brought about discussion on whether olfactory testing has enough specificity to be a useful biomarker (Wesson, Wilson, & Nixon, 2010). Olfactory function is variable within a population: over half of adults over 60 have olfactory problems, but not all have AD, suggesting that olfactory testing alone cannot differentiate health conditions. If an individual experiences a decline in olfaction, the decline could be due to a number of issues and may not accurately indicate that the individual is developing AD as opposed to some other neurodegenerative disease. It is reasonable to assume that specificity would be an issue when detecting a potential conversion to neurodegenerative disease 10-20 years before symptoms develop: pathology could develop in any number of ways because it is partially dependent on multiple environmental and genetic factors that influence disease progression along the life course. Single measurements are unlikely to ever differentiate

disease diagnoses accurately, however, the analysis of disease trajectories over time may be able to accomplish this.

Many of the olfactory tests currently available exhibit ceiling effects, and finding a way to differentiate between high performers in normative samples using an adaptive test would be beneficial in detecting smaller losses in smell in healthier (and younger) populations. Environmental conditions (temperature, humidity) as well as the participants' health (i.e., cold, sinus infection, medications) can contribute to test reliability, so olfactory testing should ideally be conducted using a burst design at each wave of measurement, or with frequent intervals, to provide a clear picture of individual olfactory health.

The above mentioned studies and criticisms highlight how important it is to compare previous level of functioning to current level of functioning in the same person. Intraindividual differences could provide the most promising information to detect changes in an individual's level of cognitive functioning and provide insight on disease trajectories.

Current Research on Early Detection of Alzheimer's Disease

A biomarker is a measurable substance or effect that is used to expedite the early detection and diagnosis of diseases. Effective biomarkers are accurate, cost effective, and easy to translate for general use. One example of a well-known biomarker is blood pressure for the early detection of heart disease. Blood pressure is regularly monitored throughout the lifespan in order to catch hypertension (persistently high blood pressure). Hypertension is the precursor to heart disease, preceding heart attacks, stroke, and cardiovascular disease. Therefore, medication to manage hypertension is given prior to heart disease in order to mitigate damage to the heart muscle and avoid the more serious heart conditions that are known to follow when left untreated. Once an individual has a heart attack the heart muscle becomes too weak and damaged for treatments to

be effective when compared to early treatment. Thus, an effective biomarker (blood pressure) provides early detection and allows for efficacious treatments to be implemented.

Biomarkers.

For Alzheimer's disease, three main clinical biomarkers are currently being developed: (a) cerebrospinal fluid (CF) protein levels (*p*-tau and amyloid); (b) neuroimaging (such as functional MRI), which is the most costly, labour intensive biomarker, and the only one currently in distributed clinical use; (c) blood sampling, currently not in use, and the most challenging to develop due to the blood brain barrier (Fiandaca et al., 2014). Though fluctuating accuracies (from 60% - 90%) have kept blood-sampling from use, it seems to be a viable alternative to the more costly neuroimaging that is currently used (Richens et al., 2014). A brand new, and promising, biomarker is retinal imaging (Frost et al., 2013). This biomarker would be measured by an ophthalmologist using specialized equipment and would likely be more cost effective than the other three in current development. A specialized machine is used to look at the retina after the patient consumes curcumin, a concentrated version of turmeric spice that binds to amyloid plaques. Doctors can look through the pupil and see the fluoresced plaques with their own eye (Frost et al., 2013). However, this technology is a long way from being implemented and is also dependent on the healthcare system, specialized equipment, and trained professionals.

Jack et al. (2010, 2013) have proposed a model of biomarkers for AD that categorizes all available biomarkers into two main categories: measuring either Amyloid-β (Aβ) deposits (CF levels, PET scans) or neurodegeneration (fMRI, t-tau levels; Jack Jr et al., 2013). This model attempts to temporally order biomarkers as they cross the threshold into abnormality in relation to each other and to disease progression (Jack Jr et al., 2013). AD pathology is likely complicated by co-morbidity with other disease pathologies, but the sequence of temporal

ordering indicates that after plaques and tangles develop there is a time lag before cognitive decline becomes behaviourally apparent. This is the key space where olfactory testing can be of most use: before the appearance of cognitive decline, at the point where pathology is beginning to develop.

Many of the methods currently in use are invasive, expensive, require healthcare professionals, or some combination of all three. An alternative to these methods is measuring olfactory performance in the population. Olfactory tests can be performed easily without a professional, are non-invasive, and have minimal cost (i.e. ~$15 vs a $3,400 PET scan). With further research, a regular olfactory screening test could detect individuals who are at risk for neurodegenerative diseases, collection of this data could then provide information on health and cognition throughout this preclinical stage, and facilitate recruitment of high risk individuals into clinical trials.

Neurobiology of Smell.

The olfactory system, unlike other sensory systems in the body (which have thalamocortical circuit relays), connects directly into the cortical regions of the brain (Purves et al., 2001). The olfactory receptor neurons are located in the olfactory epithelium inside the nasal cavity. Each olfactory receptor neuron projects onto the olfactory bulb's surface which is made up of modules called glomeruli that are comprised of glob-like tangles of axons from the olfactory receptors (Masurkar & Devanand, 2014; Shepherd, 2007). The olfactory receptor neurons synapse with termination in the glomeruli, and the resulting pattern of activity indicates a unique odour, an "odour map", which can be seen with fMRI (Purves et al., 2001; Shepherd, 2007). The fMRI indicates whether the receptors are receiving a signal. In cases where there is damage to the receptors, a signal would not be present and the individual would no longer be able to smell.

Presence of a signal indicates that the receptors are reciving a signal and an individual is most likely able to detect a smell. However, if they cannot *identify* a known smell, higher processing of smell in other brain areas may be affected (as in AD). After the olfactory receptor neurons, the odour signal is processed in the olfactory cortex (collectively comprised of the entorhinal cortex, piriform cortex, olfactory tubercle, and the amygdala) and is then projected to the orbitofrontal cortex (conscious perception of smell), as well as the thalamus and hypothalamus (motivational aspects of smell; Purves et al., 2001; Shepherd, 2007). Within this pathway, the entorhinal cortex is thought to also connect with the hippocampus, an area which contributes to odour memory (Purves et al., 2001) as well as other forms of memory. The cortical olfactory areas have multiple complex relations with limbic areas involved in learning, memory, motivation, and emotion (Shepherd, 2007). Individuals who can no longer recall the names of scents they are familiar with (odour identification test) have damage in these brain areas, affecting higher level processing of smell but not the ability to detect the smell. The sense of smell is the least understood of all the senses and many of the details involved in the olfactory perception pathways have only just begun to be explored in detail. As of yet, the circuitry and the details of synaptic organization of the orbitofrontal cortex are still unknown, as is the extent of the olfactory receptors' functionality (Shepherd, 2007).

Olfactory receptors are a subset of G protein-coupled chemical receptors found on the surface of specialized cells allowing them to sense what is going on around them (Buck & Axel, 1991; Fiandaca et al., 2015). They are called 'olfactory receptors' only because they were first discovered in the nose, but they have recently been found in lung tissue (Gu et al., 2014), spermatozoa (Eisenbach & Giojalas, 2006; Spehr et al., 2003), and in nearly every organ in the human body (Feldmesser et al., 2006). Presumably, their great abundance in skin allows for other

forms of chemical detection, such as spermatozoa finding their way to eggs (Eisenbach & Giojalas, 2006), as well as the more commonly known: odour detection through the nasal cavity, smell (Bockaert & Pin, 1999). These other functions for 'olfactory receptors' are only just beginning to be explored.

Olfaction has critical functions in health an behaviour (Croy et al., 2014). The olfactory system (olfactory epithelium, olfactory bulb, and hippocampus) have the ability to regenerate and the loss of this ability tends to signal a deeper underlying issue (Shepherd, 2007). This has led many researchers to explore loss of smell in a variety of neurodegenerative diseases including: Alzheimer's disease, schizophrenia, Huntington's disease, and idiopathic Parkinson's. Of the three testable olfactory abilities (threshold, differentiation, and identification), a loss in smell identification, the ability to smell, coupled with the loss of ability to correctly identify smells often indicates damage in the olfactory areas of the brain, episodic memory loss, and cognitive impairment (Sun, Raji, MacEachern, & Burke, 2012). The olfactory tests are not currently differentiating between a processing (i.e., interpretation) error and a memory (i.e., remembering) error. For example, if an individual is perceiving the smell of banana as pineapple they would get that wrong on the test just the same as an individual who could not remember that smell is called "banana". This could be ameliorated by adding "I don't know" and "I can't smell it" to the tests.

Alzheimer's Disease and Smell.

Olfactory deficits in patients already diagnosed with Alzheimer's disease are well known (Devanand et al., 2000; Doty et al., 1987; Fiandaca et al., 2014; Kovács, 2004; Pinto et al., 2014; Wilson et al., 2009). However, it is becoming clear that preclinical dementia (the presence of brain damage before symptoms have occurred) can be identified using olfactory testing (Albers et al., 2006; Peters et al., 2014; Wilson et al., 2009). In 2009, Wilson et al. discovered that a

person who made four errors on an olfactory test was approximately 50% more likely to develop some form of mild cognitive impairment (MCI) compared with a person who made only one error. This corroborates similar findings by Devanand et al. (2000), who found that patients who received low scores on olfactory testing, and were also unaware that their sense of smell was poor, were more likely to develop Alzheimer's than other patients. The inability of individuals to determine that their own sense of smell is declining has been replicated in many studies (Nordin, Monsch, & Murphy, 1995; Murphy et al., 2002; Croy et al., 2014). This means that someone at risk of developing Alzheimer's will not only have a declining sense of smell, but they will also be completely unaware of it, reinforcing a need for regular testing.

More recently, Peters et al. (2014) compared olfactory event-related potentials (ERPs) when exposed to scents in the brains of healthy individuals (N=8) and those diagnosed with Alzheimer's disease (N=14) or mild cognitive impairment (N=8; Peters et al., 2014). Peters et al. (2014) used Sniffin' Sticks odour pens to assess odour identification, odour threshold, and odour detection, the results of the three subtests are summed to create a composite score (TDI). A TDI score of 15 or less, out of 48, indicates anosmia. There was a large, 8 point, difference in TDI between normal individuals (32.71) and those with MCI (25.41), with a very small, 1 point, difference between MCI (25.41) and Alzheimer's (23.96; Peters et al., 2014). Additionally, Peters et al. (2014) reports that the ERP results confirmed the TDI scores, meaning anosmics (low TDI scores) were also more likely to be missing the corresponding ERP signal for olfaction (i.e., assuming the signal is present at the receptors, it is not making it to the brain). There appears to be a clear decrease in olfactory performance from normal to MCI to AD, with the largest decrease visible between normal and MCI. Although the Peters et al. (2014) sample size was small, it corroborates larger studies which have found similar results and supports the

olfactory testing evidence with ERP data, further demonstrating that olfaction may be a good marker for the beginning of cognitive impairment (Djordjevic, Jones-Gotman, De Sousa, & Chertkow, 2008; Masurkar & Devanand, 2014; Peters et al., 2014).

However, does this also mean that high cognitive functioning will lead to a better sense of smell in healthy individuals? A study to determine the relative contributions of demographic and cognitive factors to three different olfactory tasks found a significant influence of cognition on odour discrimination and identification, but not on threshold (Hedner, Larsson, Arnold, Zucco, & Hummel, 2010). Participants who performed well in executive functioning (digit span backward test) also discriminated and identified more odours correctly (Hedner et al., 2010). In as much as the digit span backward test represents a good marker for executive functioning, the findings by Hedner et al. (2010) suggest that executive function is a significant factor in discrimination and identification performance (Hedner et al., 2010). This study, as well as several previous ones, suggest that level of cognition has a significant influence on higher order olfactory performance (Hedner et al., 2010; Dulay et al., 2008; Larsson et al., 2004).

Literature Review: Longitudinal Studies

Odour Identification and Cognition.

Very few longitudinal studies have explored the use of olfactory decline as a biomarker for neurodegenerative disease (Albers et al., 2006; Doty & Kamath, 2014; Sun et al., 2012). However, papers that have examined the relationship between odour identification and disease strongly support using olfactory decline as a preclinical biomarker for detecting the risk of developing neurodegenerative disease very early on in the disease trajectory (Wilson et al., 2006; Wilson et al., 2009; Devanand et al., 2015; Devanand et al., 2000).

In 2006, Wilson et al. reported several mixed models looking at the relationship of odour identification performance to baseline level and annual rate of change in five different cognitive domains including, perceptual speed, episodic memory, semantic memory, working memory, and visuospatial ability using data from the Rush Memory and Aging Project (MAP). At this time there were a maximum of 3 waves of BSIT collection (mean of 2.7 annual evaluations per individual) and 481 participants were included in the sample for analyses (Wilson et al., 2006). Of the five domains, only perceptual speed and episodic memory were related to both baseline level of function and rate of decline, although all domains were significantly related to baseline level of function (Wilson et al., 2006). After including terms to control for age, sex, and education associations with baseline cognition and rate of cognitive change, lower odour identification scores were associated with a more rapid decline in perceptual speed (parameter estimate = 0.015, SE = 0.007, $p < 0.05$) and episodic memory (parameter estimate = 0.012, SE = 0.006, $p < 0.05$; Wilson et al., 2006). Current theory suggests that the buildup and progression of AD pathology in specific regions of the brain affects both the cognitive domains and the olfactory abilities that are associated with those areas. These results support the corresponding hypothesis that smell and episodic memory are associated. Damage to the olfactory areas of the brain indicate a higher probability of further brain damage therefore people with damage to these areas are more likely to display declines in both.

Olfactory Decline Predicts Alzheimer's Disease.

Again using data from a subgroup of Rush Memory and Aging Project (MAP) participants, Wilson et al. (2009) studied 471 individuals with valid smell (BSIT) and cognitive functioning scores. Of these, 79.3% were women and the average annual follow-up was 5.5 years. A proportional hazards model to test the relationship between odour identification score and the

risk of mild cognitive impairment (MCI), controlling for age, sex, education, presence of the ε4 allele, and level of episodic memory function, suggested that risk of MCI was associated with odour identification test scores (hazard ratio = 0.874; 95% CI: 0.812, 0.941). A person who made four errors on the odour identification test (score = 8) was 50% more likely to develop MCI than a person who made 1 error (score = 11). To test the relationship of odour identification to change in episodic memory, they used a mixed-effects model with a composite episodic memory measure as the outcome. While accounting for sex, age, education, and ε4 allele, they found that lower identification scores were associated with a faster decline in episodic memory (parameter estimate = 0.014, SE = 0.004, $p < 0.001$). The composite measure for episodic memory was created by converting raw scores of seven tests (immediate and delayed recall of the East Boston Story and Story A from Logical Memory and Word List Memory, Word List Recall, Word List Recognition) to z-scores, using the mean and standard deviation from the full cohort, then averaging the z-scores to yield the composite. Lastly, using a sub-sample of 34 individuals who had no evidence of cognitive impairment before death, they regressed a composite measure of Alzheimer's disease pathology on odour identification score in a linear regression model adjusted for age at death, sex, education, time from olfactory test to death, ε4 and composite measure of episodic memory. The 34 individuals underwent brain autopsy at death and the crude accounts of neuritic plaques, diffuse plaques, and neurofibrillary tangles in 5 brain regions were converted to a standard scale then averaged (the resulting distribution was skewed and they used a square root transformation on the data before analysis). Wilson et al. (2009) found that a lower odour identification score was associated with a higher level of Alzheimer's Disease pathology (parameter estimate = -0.063, SE= -0.027, p = 0.028). Previous work in the same data by Wilson, Arnold, Schneider, Tang and Bennett (2007) found a slightly stronger association between

neurofibrillary tangles and BSIT scores (estimated coefficient = -1.58, SE = 0.45, $p < 0.001$) versus neuritic plaques (estimated coefficient = -1.01, SE = 0.43, $p = 0.021$), with no association between BSIT and diffuse plaques. These are some of the most promising publications for advocating the use of olfactory decline as a biomarker for AD as they linked an increase in the square root of the number of neurofibrillary tangles and neuritic plaques to a lower score on an olfaction test. Using autopsy diagnoses is rare and most studies only have "potential AD" diagnoses. The autopsy sample was fairly small, but corroborates what other studies have found, and the results should encourage further studies. Wilson et al. (2009) did not adjust for hippocampal volume, which may have some effect on the results as hippocampal volume may mediate cognitive decline (Erten-Lyons et al., 2013).

Olfaction Predicts Transition to Alzheimer's Disease.

Devanand et al. (2014), using a cognitively intact community cohort of North Manhattan residents aged 65 and older, examined whether olfactory decline could predict cognitive decline in older adults without dementia. A neuropsychological test battery was administered to each participant as well as the UPSIT (the 40-item olfactory test on which the BSIT was based) to test olfactory function. Transition rates to dementia at final follow up for participants with a baseline status of no MCI was 7.03% (35/498), with 24.81% (32/129) of non-amnestic MCI at baseline transitioning to dementia, and 32.3% (42/130) of amnestic MCI transitioning to dementia. Similar results were found for the transition to Alzheimer's disease at final follow-up: no MCI at baseline 6.64%, non-amnestic MCI 24.03%, amnestic MCI 28.46%. Participants with no MCI at baseline and who later developed cognitive decline (13.67%) had a mean UPSIT score of 24.28 (SD 6.35) compared to 28.54 (SD 6.2) for those with no MCI at baseline and who did not experience later cognitive decline. This is a 4.26 point difference, but the study did not adjust for

brain volume (size) and this may have an impact on the results as brain volume can mediate cognitive decline (Erten-Lyons et al., 2013). Devanand et al. (2014) found low odour identification scores were strongly associated to later cognitive decline. When only UPSIT and Selective Reminding Test–total immediate recall (SRT-TR) were included in logistic regression UPSIT remained significant (RR 1.094 per point interval; 95% CI 1.057, 1.133), while SRT-TR was not significant, indicating that odour impairments were better than verbal assessments at predicting cognitive decline in healthy participants (Devanand et al., 2015).

Devanand et al. (2015) also used discrete time interval survival analyses and found lower baseline scores were associated with the transition to AD (hazard ratio 1.072 per point interval; 95% CI 1.036, 1.109; $p < 0.0001$) after including sex, age, education, SRT-TR and functional impairment as covariates in the model. Devanand et al. (2015) only used adults 65 and older individuals who are likely to have already experienced decline, and those trajectories are not captured here. Previous studies have found that olfactory decline due to aging begins as early as 50 years of age (Attems et al., 2015; Murphy et al., 2002) and studies would need to include participants even younger than this to capture a true baseline score and to better evaluate the decline trajectories with disease.

Current Study

Mixed effects models examining differences in the rate of change in olfaction and cognitive function between AD and non-AD individuals help to determine the potential viability of olfactory decline as a biomarker for Alzheimer's Disease (AD).

Aim 1. The Memory and Aging Project (MAP) now has more than double the data (Table 1) since Wilson et al. (2009) published results on the link between olfactory decline and cognitive

decline. The current study reproduces the methods and subsetting criteria in this more mature version of the data.

Aim 2. The focus is on solidifying the hypotheses that declining olfaction is linked to underlying AD pathology by: (a) adding autopsy diagnosed Alzheimer's disease pathology into the reproduced Wilson et al. (2009) model; and (b) conducting a separate linear mixed model with BSIT as outcome to examine the relationship between AD pathology and BSIT score.

Aim 3. The final aim examines within-person coupled variation in episodic memory and olfaction using person-mean centered BSIT. This analysis will provide answers as to whether olfaction and episodic memory are fluctuating together at each measurement time point. Using a coupling model, specifically a growth model for episodic memory with BSIT as time-varying covariate, we can look at the variability in both olfaction and episodic memory scores and determine whether assessment-to-assessment variation in olfaction mirrors variation in cognition for any given individual over time.

Table 1. Characteristics of Cognitively Normal Participants at Baseline by BSIT Group at Baseline Current Study and Wilson Study.

	Current Study (N = 1089)	Wilson Study, 2009 (N = 471)	Wilson Study, 2006 (N = 481)
Age, years (SD; range)	78.36 (7.59)	79.3 (7.0; 55-100)	80.6 (7)
Education, mean (SD; range)	14.67 (3.2; 3-3)	14.6 (2.9; 5-28)	14.5 (3)
Sex, Female%	77%	76%	72.6%
White, non-Hispanic, %	93.6%	92%	94.6%
APOE ε4 allele, %	20.1%	20%	-
Episodic memory (SD)	0.32 (0.5)	0.39 (0.5)	0.111 (0.66)
BSIT, mean (SD)	9.44 (1.95)	9.2 (1.9)	8.8 (2.2)
Waves, maximum	18	5.5	3

Methods

Participants

The current study uses two subsets of MAP data: (1) the 'reproduction subset', which follows criteria from Wilson et al. (2009) with a key aspect being the exclusion of participants with cognitive impairment at baseline; and (2) the 'extension subset', which does not exclude those with cognitive impairment at baseline. All participants were recruited from more than 30 residential facilities across the Chicago metropolitan area and were assessed annually for up to 18 years.

The reproduction subset (Table 2), consisting of 1089 individuals, was used for reproduction of Wilson et al. (2009, 2006) results (Models EP0, EP0b, EP1; Tables 4 and 5). Subset participants were selected based on having both a valid Brief Smell Identification Test (BSIT) score and clinically assessed as cognitively normal at baseline. At baseline, mean age was 78.71 years, 76% of participants were women, 21% had apolipoprotein ε4 allele, and approximately 6% were members of a racial or ethnic minority group. Over all waves, 5% of participants received a diagnosis of dementia, 11% had a stroke, and, of the 33% with autopsy data, 65.4% had intermediate to high levels of AD brain pathology, and 8.6% had a clinical diagnosis of dementia (85% of those with a clinical dementia diagnosis also had AD pathology at death).

Table 2. Reproduction Subset: Characteristics of Cognitively Normal Participants by BSIT Group at Baseline.

	All BSIT: 1-12 (N = 1089)	**Normal** BSIT: 11-12 (n = 359)	**Hyposmic** BSIT: 6-10 (n = 682)	**Anosmic** BSIT: 0-5 (n = 48)	***P* value**
Sex, Female, %	76%	80%	77%	65%	**< 0.001**
Age, years, mean (SD)	78.71 (7.4)	76.7 (7.6)	79.07 (7.6)	82.62 (5.3)	**< 0.001**
APOE ε4 allele, %	21%	21%	20%	20%	**0.043**
Cognitive domain, memory	-	-	-	-	-
Episodic, z-score (SD)	0.34 (0.5)	0.47 (0.5)	0.31 (0.5)	0.09 (0.5)	**< 0.001**
Global, z-score (SD)	0.26 (0.45)	0.37 (0.5)	0.22 (0.45)	0.00 (0.4)	**< 0.001**
Cognitive test scores	-	-	-	-	-
MMSE, mean (SD)	28.45 (1.6)	28.73 (1.4)	28.42 (1.6)	27.5 (2.3)	**< 0.001**
Logical Memory IIa, mean (SD)	10.45 (3.8)	10.97 (3.9)	10.29 (3.8)	9.04 (3.7)	**< 0.001**
Final dementia diagnosis	-	-	-	-	**< 0.001**
Normal, % (ε4 allele)	19% (14%)	11% (6%)	19% (16%)	13% (0)	
MCI, % (ε4 allele)	55% (48%)	15% (9%)	55% (27%)	23% (0)	
AD, % (ε4 allele)	71% (74%)	35% (32%)	70% (82%)	48% (25%)	
Education, years, mean (SD)	14.66 (3.3)	14.78 (3.4)	14.53 (3.3)	15.1 (4.5)	**0.006**
AD pathology, N (%)	237 (59%)	29 (51%)	124 (58%)	15 (65%)	0.092

Note: Chi-squared test for categorical or Kruskal Wallis test for continuous variables.

The extension subset (Table 3) is used in all models that include AD pathology data and consists of 1843 participants (Models EP2, EP3, and BSIT1; Tables 5 and 6). Participants mean age at baseline was 79.94, 74% of participants were women, 24% had one or more apolipoprotein ε4 allele, and approximately 6% were members of a racial or ethnic minority group. Over all waves, 11% of participants had a diagnosis of dementia, 13% had a stroke, and, of the 37% with autopsy data, 69% had intermediate to high levels of AD pathology present in the brain, and 11% had a clinical diagnosis of dementia (83% of those with a clinical dementia diagnosis also had AD pathology at death). This subset was used for Models EP2, EP3, and BSIT1 because once AD pathology is included in the model it is no longer relevant whether individuals were clinically assessed to be cognitively normal at baseline: pathology may be occurring in those individuals as well.

Table 3. Extension Subset: Characteristics of Participants by BSIT Group at Baseline.

	All BSIT 1-12 (N = 1852)	**Normal** BSIT 11-12 (n = 423)	**Hyposmic** BSIT 6-10 (n = 965)	**Anosmic** BSIT 0-5 (n = 131)	***P* value**
Sex, Female%	74%	79%	75%	61%	< 0.001
Age, years, mean (SD)	79.94 (7.6)	77.18 (7.8)	79.93 (7.7)	83.77 (6.7)	< 0.001
APOE ε4 allele, N (%)	387 (24%)	75 (21%)	187 (23%)	38 (34%)	0.002
Cognitive domain	-	-	-	-	-
Episodic memory, z-score (SD)	0.00 (0.8)	0.35 (0.6)	0.03 (0.7)	-0.68 (0.9)	< 0.001
Global memory, z-score (SD)	-0.01 (0.7)	0.27 (0.5)	0.02 (0.6)	-0.6 (-0.71)	< 0.001
Cognitive test scores	-	-	-	-	-
MMSE, mean (SD)	27.4 (3.5)	28.5 (1.7)	27.74 (2.4)	24.95 (4.3)	< 0.001
Logical Memory IIa, mean (SD)	8.73 (4.7)	10.34 (4.1)	8.8 (4.5)	5.62 (4.6)	< 0.001
Education, years, mean (SD)	14.61 (3.3)	14.73 (3.4)	14.5 (3.2)	14.73 (3.8)	< 0.001
AD pathology, N (%)	438 (65%)	39 (53%)	204 (63%)	48 (73%)	< 0.001

Measures

The Brief Smell Identification Test.

The Brief Smell Identification Test (BSIT) was used to assess odour identification. The BSIT is the 12-item version (Doty, Marcus, & Lee, 1996) of the 40-item University of Pennsylvania Smell Identification Test (UPSIT; Doty, Shaman, & Dann, 1984; Doty, Shaman, Kimmelman, & Dann, 1984). The test is in a booklet format with one encapsulated scent and 4 answer choices per page. Each scent microcapsule contains a culturally familiar odour which is scratched with a pencil and placed under the nose of the participant. The participant then chooses one of the four options provided in the booklet and the score is the number of odours correctly recognized. Keeping with standard practice, missing responses are assigned a score of 0.25, to a maximum of two missing answers. If more than two responses are missing, the entire test is treated as missing, as described in previous works (Wilson et al., 2007). Performance on the 12-item BSIT has been shown to correspond to the 40-item UPSIT from which it was derived (Makowska, Kloszewska, Grabowska, Szatkowska, & Rymarczyk, 2011), and has a test-retest value of $r = 0.71$ (Doty et al., 1996). Cronbach's alpha, a measure of internal consistency, represents how closely related a set of items are as a group. It is not a statistical test, but a coefficient of reliability and can be written as a function of the number of test items and the average inter-correlation among the items, with 0.7 or higher being acceptable by most textbook standards. In MAP data, the BSIT Cronbach's coefficient alpha was 0.68 (Wilson, Yu, & Bennett, 2010). MAP participants consist of mainly white individuals and other studies have found varying degrees of internal consistency with the BSIT. For example, in a multiethnic cohort, Menon et al. (2013)

found that internal consistency was moderate (Cronbach's α coefficient .60), which is similar to other multiethnic studies (Hawkins et al., 2011). In contrast, the full UPSIT has an Internal Consistency Rating (calculated from the split-half correlations and the Spearman-Brown formula) of 0.92 (Doty, Frye, & Agrawal, 1989; Doty et al., 1996).

Another limitation to consider when using the BSIT over the UPSIT is that the norms included in the BSIT manual are extrapolated from UPSIT norms (Doty, 2001) even though many of the distractor smells have been changed in the BSIT compared to the UPSIT (Menon, Westervelt, Jahn, Dressel, & O'Bryant, 2013).

Most studies agree that the number of years of education does not significantly affect BSIT performance, supporting claims that BSIT is relatively resistant to sociodemographic factors (Doty et al., 1996; Goudsmit et al., 2003; Hawkins & Pearlson, 2011; Menon et al., 2013). However, the sensitivity of the BSIT, which has ranged from 67%-79% (Kjelvik, Sando, Aasly, Engedal, & White, 2007; Menon et al., 2013) could be improved. It is important to keep in mind that because the test is forced choice, with only 4 answer options, even anosmic responders will potentially guess the correct answer and achieve an average of 25%. Increased information and sensitivity could be attained by adding alternatives to the answer choices such as: 'I don't know' or 'I can't smell it'.

Episodic Memory Composite.

Raw scores from seven tests (East Boston immediate recall, East Boston delayed recall, Logical memory I (immediate recall), Logical memory II (delayed recall), CERAD word list, (immediate (three trials), delayed and recognition) were converted to z-scores and averaged to yield an episodic memory composite score. The z-score describes how much

a score deviates from the mean, for example, a negative z-score indicates that someone has an overall score that is lower than the average of the entire cohort at baseline.

East Boston Delayed and Immediate Recall.

After a three sentence story is read, participants are asked to retell it from memory immediately (East Boston Immediate Recall) and after a 3-minute distraction (East Boston Delayed). Both are scored as the number of story units recalled (out of 12).

Logical Memory I (immediate recall) and II (delayed recall).

Logical Memory is a measure from the Wechsler Memory Scale, revised in 1987, where a brief story is read to the participants. Participants are asked to retell it from memory immediately (I) and after an approximately 30-minute delay (II). Both are scored as the number of story units recalled (out of 25).

CERAD Word List immediate (3-trial), delayed, and recognition.

Word list immediate (3 trial) is a measure from the CERAD set of neuropsychological assessment battery. A 10-word list is presented three times for a total of 30 words, with three immediate recall trials. The total score is out of 30, the number of words recalled over all 3 trials.

In word list delayed, participants are asked to recall the same ten words that were presented in the previous immediate recall measure after a several minute delay. The number of words recalled out of 10 is scored.

In word list recognition, ten sets of 4 words are presented. Participants select the word from each set that they recognize from the previous trials (immediate and delayed) among the distractor words. The test is scored out of 10.

Global Cognitive Function Composite.

This is a composite of 19 tests, including all the episodic memory composite tests plus the following: Boston naming, category fluency, reading test, digits forwards, backwards and ordering, line orientation, progressive matrices, symbol digits modality, oral number comparison, Stroop colour naming, and Stroop word reading. As in the episodic memory composite measure, raw scores from the 19 tests were converted to z-scores and averaged to yield a global cognitive summary score. Domains covered by this composite include: episodic memory, working memory, semantic memory, perceptual speed and perceptual orientation.

Alzheimer's Disease Pathology.

A binary score for AD pathology was used in all analyses. This measure relies on both neurofibrillary tangles and neuritic plaques. A neuropathological evaluation was done without knowledge of clinical information, including a diagnosis of dementia. The neuropathologist then determined the level of AD pathology on a 4-point scale. Those with intermediate or high scores fulfill criteria for having a pathologic diagnosis of AD. This measure was coded as: 1 = high AD pathology; 2 = intermediate AD pathology; 3 = low AD Pathology and 4 = no AD pathology. For the current analyses, these were grouped into a binary variable which was coded as 1 or 2 = 1; 3 or 4 = 0.

Clinical Dementia Diagnosis.

Clinical dementia diagnosis is measured at every assessment and is based on: (1) computer scoring of 19 cognitive tests that cover 5 cognitive domains (see Global Cognition); (2) clinical judgment by a neuropsychologist; and (3) diagnostic classification by a clinician (neurologist, geriatrician, or geriatric nurse practitioner) who

then reviews all available data, examines the participant, and gives a final diagnostic classification.

The resulting variable is assigned a value from 1 to 6, and includes: (1) no cognitive impairment, (2) MCI, (3) MCI plus another condition affecting cognition, (4) AD, (5) AD plus another condition affecting cognition, and (6) other dementia. This variable was used to exclude individuals who were clinically diagnosed with cognitive impairment at baseline in the reproduction subset (i.e., only those with a 1 at baseline were included).

BSIT_PM.

The person mean BSIT variable (BSIT_PM) is each individual's own personal mean for BSIT score. It appears by itself as a level 2 predictor in order to account for whether people with better BSIT (overall) score better on the cognitive test. It also appears as an interaction with time (timeinStudy*BSIT_PM) to account for whether people with better BSIT decline more slowly over time. In individuals whose BSIT is declining over time, and at varying rates, this variable will partially capture this because the later, lower, scores will bring down those individuals' average.

BSIT_TVC.

BSIT as a time varying co-variate (BSIT_TVC) and is coded as BSIT score at each occasion minus each person's own individual person mean score (BSIT_PM). BSIT_TVC is a fixed, within-person level-1 variable, and along with time acts as an index of how a person's score changes over time.

Apolipoprotein ε4 allele (APOE ε4).

The APOE allele variants were recorded and a new binary variable was created which grouped those who had one or more ε4 alleles into group 1 and those without an ε4 allele into group 0.

Statistical Plan

Aim 1: Reproduction.

In 2009, Wilson et al. used a proportional hazards model to test the relation of odour identification score to risk of incident MCI. The current study reproduces this model, with similar subset criteria (reproduction subset), in the same dataset but with more participants and up to 13 more years of data collection. Eliason (1993) suggests a minimum sample size of 60 is required for these models if there are 5 or fewer parameters for covariates (including treatment) to be estimated, so the available sample and model met these criteria. Different sample sizes among treatment groups pose no special difficulty (Eliason, 1993). The proportional hazards model in the current study uses the Pencina et al. (2007) suggested "entry age adjusted" time scale (Pencina, Larson, & D'Agostino, 2007). It is unclear if Wilson et al. (2009) used an entry age adjusted model or, the more commonly used, linear baseline age time-on-study model (described in Pencina et al., 2007). Lower BSIT scores are expected to be associated with poor cognitive outcomes after controlling for age, sex, education, and apolipoprotein ε4 allele (Devanand et al., 2015; Wilson et al., 2009). Cox survival analysis assumes that the shape of the survival function over time is the same for all cases and, as an extension, for all groups. Therefore, to look at differences in change over time for different groups, mixed effects models are used.

In 2006, Wilson et al. estimated five different cognitive outcome mixed effects models on a subset of MAP data with participants who had both a valid BSIT score and no dementia at baseline. This model did not include terms to control for baseline level of episodic memory. In 2009, Wilson et al. estimated a similar mixed-model using episodic memory as the outcome, but added a term to control for baseline episodic memory. The current study reproduces both of these results using similar subset criteria (reproduction subset) and a similar model: episodic memory composite is the outcome, time in study (age at visit - age at baseline) is the chronological metric, and sex, education, age, ε4 allele, baseline episodic memory (excluded for replicating 2006 model, Model EP0b), and baseline BSIT are covariates of the intercept and rate of change (Model EP0, current study; N = 927).

(Model EP0)

Level 1: $\text{Episodic Memory}_{ij} = \beta_{0i} + \beta_{1i}\,(\text{Time}_{ij}) + e_{ij}$

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(\text{AgeBL-78}) + \gamma_{02}(\text{Sex}) + \gamma_{03}(\text{Educ-14}) + \gamma_{04}(\varepsilon 4)$

$+ \gamma_{05}(\text{EpisodicMemBL-0.377}) + \gamma_{06}(\text{BSITBL-8.76}) + U_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(\text{AgeBL-78}) + \gamma_{12}(\text{Sex}) + \gamma_{13}(\text{Educ-14}) + + \gamma_{14}(\varepsilon 4)$

$+ \gamma_{15}(\text{EpisodicMemBL-0.377}) + \gamma_{16}(\text{BSITBL-8.76}) + U_{1i}$

Some would argue that episodic memory at baseline should not be included because individuals who are lower at the baseline measurement in the study could already be in cognitive decline. Peto (1981) dubbed this the “horse-racing effect”. For example, an individual whose episodic memory has been decreasing faster than average in previous

decades will now have a lower than average episodic memory score, and will also tend to have lower than average episodic memory in the future (Peto, 1981). Individuals enter a study with an entire history behind them; those who are lower at entry are less likely to "win the race". Further, the level at baseline becomes irrelevant once a later level is known, such as AD pathology. Model EP0 reproduces Wilson et al. (2009). To account for both views, Model EP0b excludes the term for baseline episodic memory. A final model, EP1, similarly does not include baseline episodic memory, and also removes the non-significant terms regressing rate of change on sex and education.

(Model EP1)

Level 1: $\text{Episodic Memory}_{ij} = \beta_{0i} + \beta_{1i}(\text{Time}_{ij}) + e_{ij}$

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(\text{AgeBL-78}) + \gamma_{02}(\text{Sex}) + \gamma_{03}(\text{Educ-14}) + \gamma_{04}(\varepsilon 4)$

$+ \gamma_{05}(\text{BSITBL-8.76}) + U_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(\text{AgeBL-78}) + \gamma_{12}(\varepsilon 4) + \gamma_{13}(\text{BSITBL -8.76}) + U_{1i}$

Here, the Level 1 equation depicts change in episodic memory performance for a given individual (i) on a given measurement occasion (j) and specifies an intercept β_{0i}, effects of linear time (β_{1i}), and the within-person residual (e_{ij}).

Level 2 specifies fixed effects for the common intercept (γ_{00}), and regression parameters for its predictors: age centered on the mean (γ_{01}), sex (γ_{02}) with female as the reference group, education centered on the mean (γ_{03}), ε4 allele (γ_{04}) with presence of ε4 allele as reference, BSIT centered on the mean (γ_{05}). Level 2 also specifies a fixed effect for linear change over time (γ_{10}), centered at each individual's first occasion of measurement, and its predictors: age (γ_{11}), ε4 (γ_{12}), and BSIT (γ_{13}). Person-specific

deviations (random effects) from the fixed intercept (U_{0i}) and fixed linear effect of time (U_{1i}) are also included.

Aim 2: Extension.

AD pathology was added to Model EP1 under the assumption that individuals with AD pathology at death would have been accumulating the pathology over time, which may then affect the rate of change in episodic memory and at baseline (Model EP2, N = 573). Since AD pathology was included as a term in Model EP2, this eliminated the need to exclude individuals with clinically assessed dementia at baseline, and the final analysis of this model was reported using the extension subset. Additionally, results from EP2 were compared using the extension subset and the reproduction subset. Results were similar, with the extension subset providing significance at $p < 0.001$ as opposed to significance at $p < 0.05$ in the reproduction subset, presumably due to the additional statistical power.

Wilson et al. (2009) also ran a linear regression, regressing AD pathology on BSIT. After adjusting for age, sex, education, time from BSIT test to death, ε4 and episodic memory, they found lower identification scores were associated with higher levels of AD pathology (Wilson et al., 2009). Instead of a regression, this paper used a mixed model with BSIT as outcome to determine whether AD pathology is related to baseline level of olfactory function and whether AD pathology may be related to the rate of change in BSIT over time (Model BSIT1, N = 574). This assumes that individuals build pathology up over an extended period of time (Jack et al. 2010), and that individuals with high pathology at death are also more likely to have entered the study with higher pathology than those who have low pathology at death. This pathology then accumulates over the course of the study and results in the high level of pathology at death. Terms for sex,

education, and ε4 on rate of change were not significant and they were not included in the final model.

(Model BSIT1)

Level 1: $BSIT_{ij} = \beta_{0i} + \beta_{1i} (Time_{ij}) + e_{ij}$

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(AgeBL\text{-}78) + \gamma_{02}(Sex) + \gamma_{03}(Educ\text{-}14) + \gamma_{05}(\varepsilon 4)$

$+ \gamma_{04}(ADpathology) + U_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(AgeBL\text{-}78) + \gamma_{12}(ADpathology) + U_{1i}$

Aim 3: Growth Model for Episodic Memory with BSIT as Time-varying Covariate.

To examine the time-varying covariation between BSIT and episodic memory (within-person coupled variation) a time in study based mixed-effects model was constructed using a composite score for episodic memory (Model EP3, N = 573). In Model EP3, the individual person mean of BSIT was used to represent a between person effect and this was subtracted from the raw score to then provide a within-person effect. To the best of my knowledge this is the first time this type of model has been used to examine the relationship between olfaction, episodic memory and AD pathology. Terms for sex and education were not significant on rate of change and were excluded in the final model.

(Model EP3)

Level 1: $Episodic\ Memory_{ij} = \beta_{0i} + \beta_{1i} (Time_{ij}) + \beta_{2i}(BSIT_{ij}\text{-}BSIT\_PM_i) + e_{ij}$

Level 2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(BSIT\_PM) + \gamma_{02}(AgeBL\text{-}78) + \gamma_{03}+(Sex) + \gamma_{04}(Educ\text{-}14)$

$+ \gamma_{05}(\varepsilon 4) + \gamma_{06}(ADpathology) + U_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(BSIT\_PM) + \gamma_{12}(AgeBL\text{-}78) + \gamma_{13}(\varepsilon 4) + \gamma_{14}(ADpathology) + U_{1i}$

$\beta_{2i} = \gamma_{20}$

Here, the Level 1 equation depicts change in episodic memory performance for a given individual ($_i$) on a given measurement occasion ($_j$) and specifies an intercept β_{0i}, effects of linear time (β_{1i}), the within-person component of BSIT (β_{2i}), the difference between occasion-specific and person-mean values), and the within-person residual (e_{ij}).

Level 2 specifies fixed effects for the common intercept (γ_{00}), the between-person component of BSIT (γ_{01}, person-mean BSIT values), fixed effect slopes for linear time (γ_{10}) and coupled within person variation between BSIT and episodic memory (γ_{20}). Person-specific deviations (random effects) from the fixed intercept (U_{0i}) and fixed linear effect of time (U_{1i}) are also included. Covariates for age at baseline (centered at the mean, 78), biological sex, APOE status (ε4 allele or no ε4 allele), and education (centered at the mean, 14) were included.

Data management was handled with R (Version 3.3.1; R Development Core Team, 2016) using R Studio (Version 1.0.136) , mixed-effects model analyses were conducted with R package lme4 (Bates, Maechler, Bolker, & Walker, 2015), graphs and tables were produced using ggplot2 (Wickham, 2009) and sjPlot (Lüdecke, 2017), respectively. Proportional hazards analyses were conducted with SPSS (Version 23).

Results

Aim 1: Reproduction

1.1 Survival Analysis.

This study found that for every additional point correctly identified on the BSIT, the risk of being diagnosed with dementia decreased by 6.4% (hazard ratio = 0.936, 95% CI: 0.699, 0.937). These results reproduced what Wilson et al. (2009) found: risk of MCI was associated with BSIT score (hazard ratio = 0.874; 95% CI: 0.812, 0.941).

The risk of dementia was greater for anosmic and hyposmic individuals compared to normosmic. Risk of dementia for those with an ε4 allele, compared to those without, was 1.5 times greater. These findings corroborate previous research in this area: low olfactory scores are related to an increased risk of dementia (Devanand et al., 2015; Wilson et al., 2009).

1.2 Mixed Effects Models EP0 & EP1.

This study found a robust association between lower BSIT scores and a more rapid decline in episodic memory in Model EP0 (parameter estimate = 0.02, SE = 0.002, $p < 0.001$), however the relationship between BSIT and episodic memory at baseline was not significant. The results from Model EP0 reproduced those from Wilson et al. (2009). After excluding baseline episodic memory, results also reproduced Wilson et al. (2006). Model EP0b found that lower BSIT was robustly associated with more rapid decline in episodic memory (parameter estimate = 0.02, SE = 0. 002, $p < 0.001$) as well as at baseline (parameter estimate = 0.04, SE = 0.008, $p < 0.001$; Table 4).

In Model EP1 there was a robust association between both lower BSIT scores and a more rapid decline in episodic memory (parameter estimate = 0.02, SE = 0.002, $p <$

0.001) and episodic memory at baseline (parameter estimate = 0.04, SE = 0.008, $p <$ 0.001). The results from Model EP1 also match Wilson et al., (2006), when baseline cognition was not a part of the model, and is considered the final model (Table 5).

Table 4. Results of Mixed Effect Reproduction Models: Episodic Memory as Outcome.

	EP0		EP0b	
	B (CI)	*p*	*B (CI)*	*p*
Fixed Effects				
(Intercept)	0.43 (0.40 – 0.45)	**<.001**	0.33 (0.29 – 0.37)	**<.001**
Time in Study	-0.04 (-0.06 – -0.03)	**<.001**	-0.04 (-0.05 – -0.03)	**<.001**
Age	-0.002 (-0.00 – 0.00)	.092	-0.01 (-0.02 – -0.01)	**<.001**
EpisodicMemoryBL	0.82 (0.78 – 0.85)	**<.001**		
Sex	-0.08 (-0.12 – -0.04)	**<.001**	-0.19 (-0.26 – -0.12)	**<.001**
Education	0.01 (0.00 – 0.01)	**.001**	0.05 (0.04 – 0.06)	**<.001**
ε4 Allele	0.06 (0.02 – 0.10)	**.003**	0.01 (-0.06 – 0.08)	.792
BSIT	0.003 (-0.01 – 0.01)	.540	0.04 (0.02 – 0.05)	**<.001**
Time:Age	-0.004 (-0.01 – -0.00)	**<.001**	-0.004(-0.00 – -0.00)	**<.001**
Time:EpisodicMemoryBL	0.01 (-0.01 – 0.03)	.297		
Time:Sex	0.01 (-0.01 – 0.03)	.296	0.01 (-0.01 – 0.03)	.403
Time:Education	0.00001(-0.00 – 0.00)	.992	0. 0006(-0.00 – 0.00)	.636
Time:ε4	-0.06 (-0.08 – -0.04)	**<.001**	-0.06 (-0.07 – -0.04)	**<.001**
Time:BSIT	0.02 (0.01 – 0.02)	**<.001**	0.02 (0.01 – 0.02)	**<.001**
Random Effects				
σ^2	0.100		0.103	
$\tau_{0,\ intercept}$	0.019		0.140	
$\tau_{1,\ slope}$	0.011		0.009	
N_{projid}	927		927	
ICC_{projid}	0.146		0.556	
Observations	6486		6486	
R^2 / Ω_0^2	.843 / .842		.850 / .848	
AIC	5965.495		7060.339	
Deviance	5931.495		7030.339	

Note. CI = confidence interval

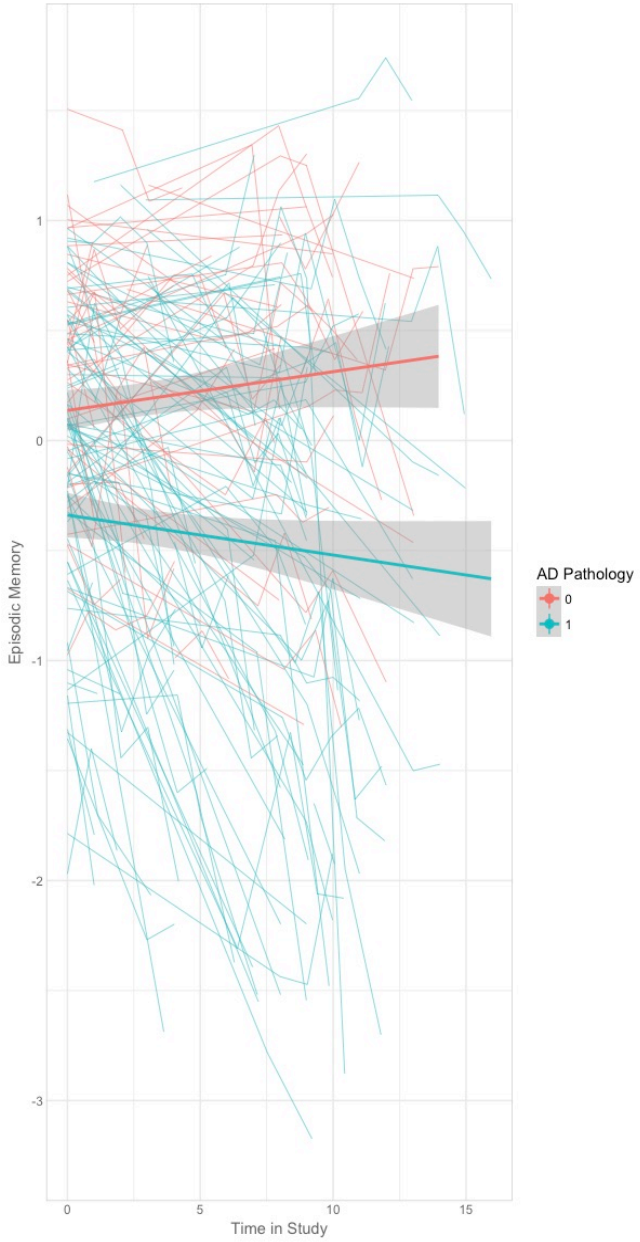

Figure 1. Raw and Average Trajectories for Episodic Memory by AD Pathology Group.

Aim 2: Extension

2.1 Model EP2: Episodic Memory + AD.

Model EP2 used the second subset of data (Table 3), where those who were clinically diagnosed with MCI or AD at baseline were not excluded, and autopsy diagnosed AD pathology was added to the model.

Both AD pathology (parameter estimate = -0.22, SE = 0.07, $p = 0.003$) and BSIT (parameter estimate = 0.12, SE = 0.02, $p < 0.001$) were significantly related to episodic memory at baseline (Table 5). This indicates that low AD pathology was related to higher episodic memory scores at baseline and higher baseline BSIT scores were related to higher baseline episodic memory scores, as expected (Table 5). Higher AD pathology (parameter estimate = -0.08, SE = 0.02, $p < 0.001$) and lower BSIT scores (parameter estimate = 0.014, SE = 0.003, $p < 0.001$) were also significantly associated with a more rapid decline in episodic memory. The average trajectories for episodic memory were visibly different for AD pathology groups (Figure 1).

Table 5. Results of Final Mixed Effect Models: Episodic Memory as Outcome.

	EP1		EP2		EP3	
	B (CI)	*p*	*B (CI)*	*p*	*B (CI)*	*p*
Fixed Effects						
(Intercept)	0.33 (0.29 – 0.37)	**<.001**	0.23 (0.09 – 0.36)	**.001**	-0.93 (-1.20 – -0.66)	**<.001**
Time in Study	-0.04 (-0.05 – -0.03)	**<.001**	-0.03 (-0.06 – -0.00)	**.034**	-0.08 (-0.15 – -0.01)	**.020**
Age	-0.01 (-0.02 – -0.01)	**<.001**	-0.02 (-0.03 – -0.01)	**.004**	-0.02 (-0.03 – -0.01)	**<.001**
Sex	-0.19 (-0.25 – -0.12)	**<.001**	-0.25 (-0.40 – -0.10)	**.001**	-0.20 (-0.34 – -0.07)	**.003**
Education	0.05 (0.04 – 0.06)	**<.001**	0.05 (0.02 – 0.07)	**<.001**	0.05 (0.03 – 0.07)	**<.001**
ε4 Allele	0.01 (-0.06 – 0.08)	.793	-0.32 (-0.48 – -0.15)	**<.001**	-0.28 (-0.43 – -0.13)	**<.001**
BSIT	0.04 (0.02 – 0.05)	**<.001**	0.12 (0.09 – 0.15)	**<.001**		
Time:Age	-0.004 (-0.01 – -0.00)	**<.001**	-0. 0004 (-0.00 – 0.00)	.758	0. 00009 (-0.00 – 0.00)	.953
Time:ε4	-0.06 (-0.07 – -0.04)	**<.001**	-0.06 (-0.10 – -0.03)	**<.001**	-0.04 (-0.08 – -0.00)	**.029**
Time:BSIT	0.02 (0.01 – 0.02)	**<.001**	0.01 (0.01 – 0.02)	**<.001**		
AD Pathology			-0.22 (-0.36 – -0.08)	**.003**	-0.24 (-0.37 – -0.10)	**<.001**
Time:AD Pathology			-0.08 (-0.11 – -0.05)	**<.001**	-0.06 (-0.09 – -0.03)	**<.001**
BSIT_PM					0.13 (0.10 – 0.16)	**<.001**
BSIT_TVC					0.07 (0.04 – 0.10)	**<.001**
Time:BSIT:PM					0.01 (0.00 – 0.02)	**.005**
Random Effects						
σ^2	0.103		0.145		0.162	
$\tau_{0,\ intercept}$	0.140		0.450		0.375	
$\tau_{1,\ slope}$	0.009		0.012		0.004	
N_{projid}	927		454		573	
ICC_{projid}	0.555		0.742		0.693	
Observations	6486		2583		909	
R^2 / Ω_0^2	.850 / .848		.891 / .889		.928 / .914	
AIC	7057.443		4076.257		1897.833	
Deviance	7031.443		4046.257		1865.833	

Note. CI = confidence interval

2.2 Model BSIT1.

In Model BSIT1, AD pathology was not related to BSIT at baseline. However, higher AD pathology at autopsy was significantly related to a faster decline in BSIT scores (parameter estimate = -0.09, SE = 0.042, $p = 0.022$; Table 6).

Table 6. Results of Multilevel Model: BSIT as Outcome.

	BSIT1	
	B (CI)	*p*
Fixed Effects		
(Intercept)	9.26 (8.85 – 9.68)	**<.001**
Time in Study	-0.08 (-0.15 – -0.02)	**.019**
Age	-0.09 (-0.12 – -0.05)	**<.001**
Sex	-0.71 (-1.15 – -0.27)	**.002**
Education	-0.01 (-0.08 – 0.06)	.786
ε4 Allele	-0.80 (-1.27 – -0.33)	**<.001**
AD Pathology	-0.38 (-0.82 – 0.06)	.091
Time:Age	-0.00 (-0.01 – 0.01)	.589
Time:AD Pathology	-0.10 (-0.18 – -0.02)	**.022**
Random Effects		
σ^2	2.117	
$\tau_{0,\ intercept}$	3.876	
$\tau_{1,\ slope}$	0.008	
N_{projid}	574	
ICC_{projid}	0.646	
Observations	914	
R^2 / Ω_0^2	.866 / .821	
AIC	4059.450	
Deviance	4035.450	

Note. CI = confidence interval

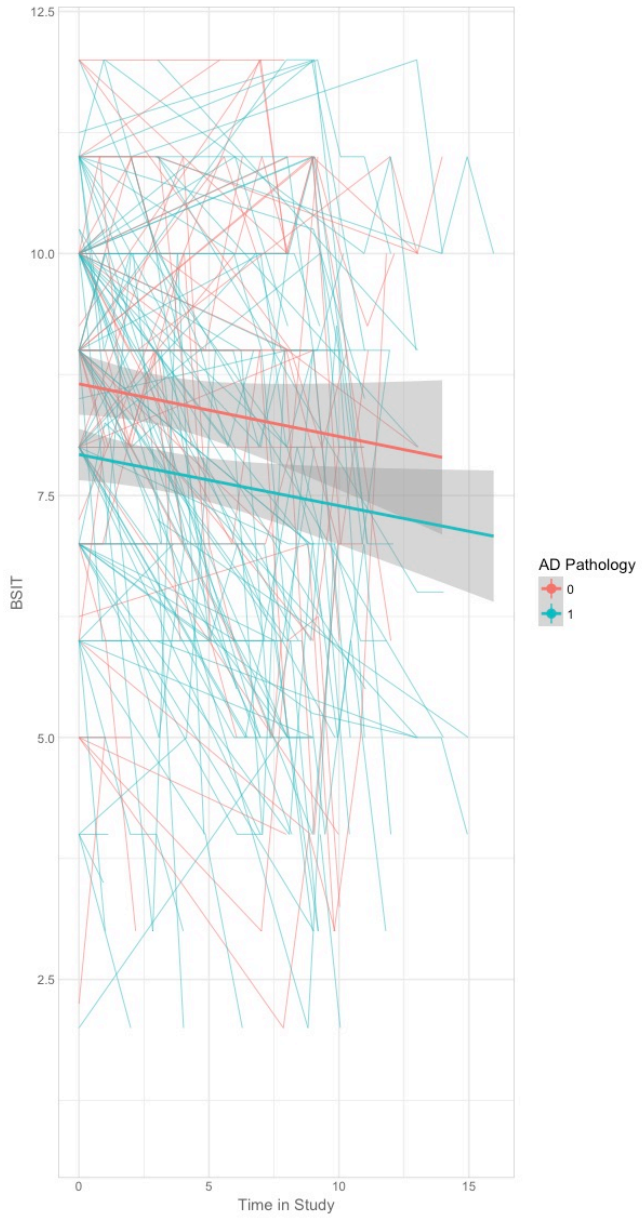

Figure 2. Raw and Average Trajectories for BSIT by AD Pathology Group.

Aim 3: Growth Model for Episodic Memory with BSIT as Time-varying Covariate

In Model EP3, the between-person variation in odour identification had a significant and robust positive association to episodic memory (parameter estimate = 0.13, SE = 0.014, $p < 001$; Table 5). For every unit more in person-mean BSIT (PM_BSIT) at baseline, episodic memory at baseline was also higher by 0.13. High AD pathology was related to lower episodic memory at baseline (parameter estimate = -0.24, SE = 0.07, $p <$ 001). Person-mean BSIT scores (parameter estimate = 0.011, SE = 0.004, $p = 0.005$) and high AD pathology (parameter estimate = -0.06, SE = 0.02, $p < 0.001$) were both associated with more rapid declines in episodic memory.

The within-person coupling effect represents the relative within-person effect of odour identification variation on variation in episodic memory. There was a robust positive association between the time varying covariate of BSIT (BSIT_TVC = raw BSIT - PM_BSIT) and fluctuations in episodic memory. For every unit increase in BSIT relative to each person's own mean BSIT score, episodic memory, for that same occasion, also increases by 0.07 units (parameter estimate = 0.071, SE = 0.02, $p < 001$).

A graph of the association between time-specific fluctuations for episodic memory and BSIT_TVC by AD pathology group (Figure 3) revealed an interaction of BSIT_TVC and AD pathology. Model EP3 was then estimated including an interaction between BSIT_TVC and AD pathology, the interaction was significant (parameter estimate = 0.075, SE = 0.034, $p = 0.03$) supporting the association seen in the graph.

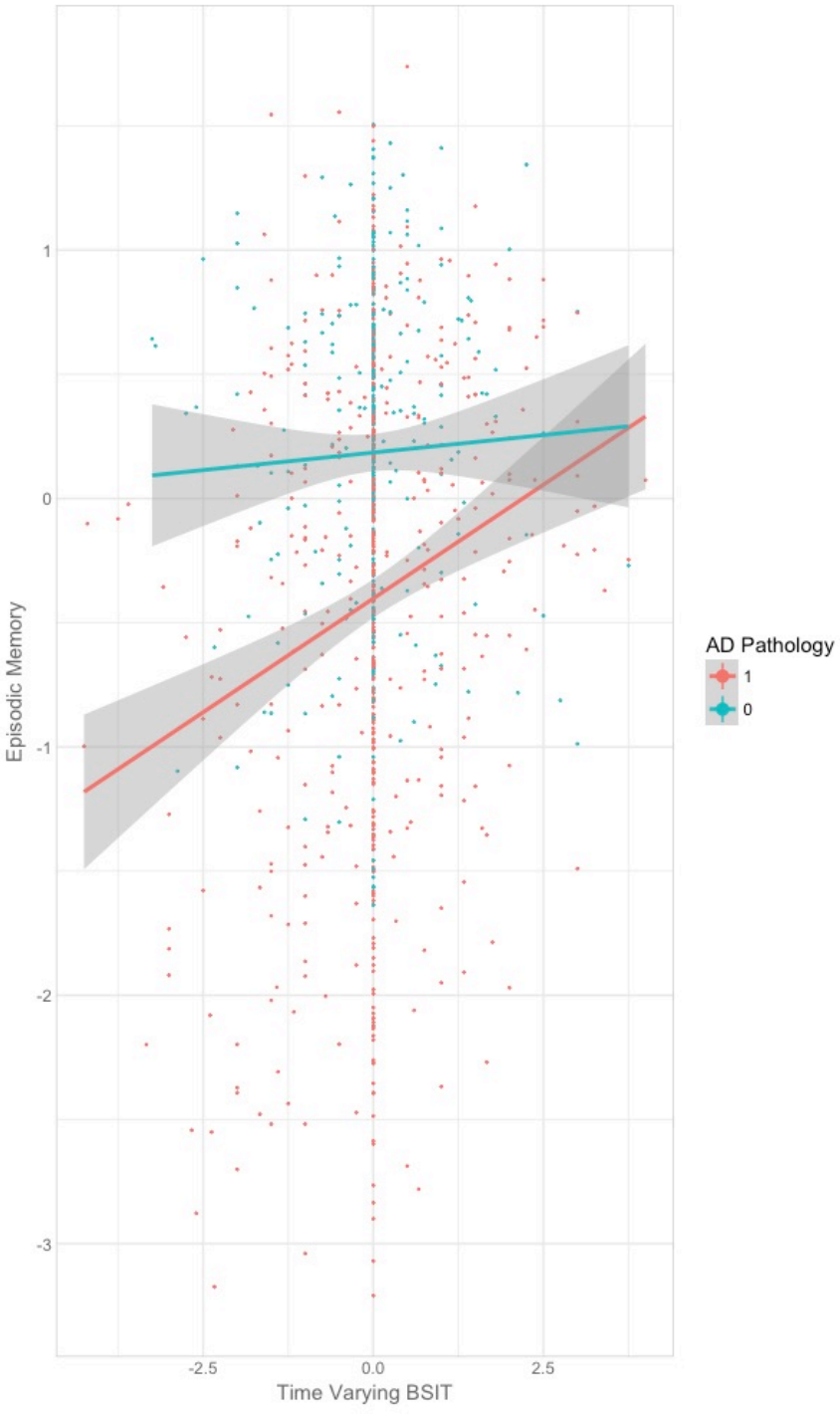

Figure 3. Association Between Time-specific Fluctuations for Episodic Memory and BSIT_TVC by Alzheimer's Pathology Group.

Discussion

Previous work has indicated that lower baseline olfactory scores are associated with lower episodic memory at baseline and faster declines in episodic memory (Wilson et al., 2006; Wilson et al., 2009). The analysis presented here, using a more mature version (more cases and more waves) of the same longitudinal study, successfully reproduced those results (Models EP0, EP0b, and EP1) as well as several interesting findings in the extension (Models EP2 & EP3).

This study's final reproduction model (EP1) did not include baseline episodic memory. In addition to the "horse-racing" effect put forth by Peto (1981), there is also evidence that when the intercept is specified at baseline the baseline episodic memory "covariate" is then predicting itself. This can be seen in EP0 (Table 4): for a 1-unit increase in baseline memory (EpisodicMemoryBL) the increase in intercept is 0.82 (95% CI: 0.78-0.85), which is very close to 1. After excluding it, results indicated that lower BSIT was robustly associated with lower episodic memory scores at baseline as well as more rapid declines in episodic memory (EP0b, EP1, EP2).

All models were estimated with uncorrelated random effects. The choice between correlating and uncorrelating random effects is based on the assumption that participants with a higher baseline than average are either likely to also have a faster rate of change than average (correlated) or not (uncorrelated or independent). Estimating the models with uncorrelated random effects can be justified due to how differences in individuals' experiences in the environment (e.g., toxins etc.) affect olfactory ability making baseline olfactory ability different for everyone, but this would not necessarily impact rate of

decline. However, in addition to this, some models would not converge with correlated random effects and AD pathology both included in the model (BSIT1 & EP3). In order to be able to compare across all models, all of them were estimated with uncorrelated random effects. Overall, results were similar for models that would estimate both correlated and uncorrelated random effects (Models EP0, EP0b, EP1, and EP2). However, comparisons between models estimated with correlated random effects indicated that there may be differences between the subsets. For example, when comparing results of Model EP2 estimated with correlated random effects in the reproduction subset versus the extension subset, there was a higher correlation in the random effects for the reproduction subset (correlation of 0.4 vs 0.2). The reproduction subset selects only those who are not experiencing any cognitive decline at baseline, whereas the models using AD pathology, the extension subset, are inherently selecting only individuals who have already died (have autopsy data). After excluding those with MCI & dementia at baseline, a homogeneous healthy population does not have much variance at baseline (e.g., using Model EP2: intercept variance component for reproduction subset is 0.1 vs 0.4 in the extension subset).

The underlying theory is that loss of odour identification reflects accumulating AD pathology in the brain that is affecting brain regions that support both olfaction and episodic memory (Ohm & Braak, 1987; Wilson et al., 2006). To test this, Model EP2 and BSIT1 (see Tables 5 and 6) were estimated. This study found that AD pathology is a significant predictor of baseline level and rate of change in memory (Model EP2), meaning that both lower BSIT scores and higher AD pathology are significantly related

to lower episodic memory scores at baseline as well as faster declines in episodic memory scores (Model EP2, Table 5). However, when BSIT is the outcome (Model BSIT1), higher AD pathology is robustly related to a faster decline in BSIT, but not to the baseline BSIT score (Table 6). This could be due in part to the small range offered by the BSIT (0-12). If AD pathology affects both BSIT and episodic memory, these findings might indicate that BSIT could be used as a proxy for AD pathology, as well as possibly explain some of the inconsistency found with BSIT at baseline (i.e., all of the variables are intertwined and the models are not accurately capturing the relationships).

AD pathology is measured only once, at death (up to 1 year after the final occasion of measurement, and 1 to 18 years after the baseline visit) and there are no variables in this data to account for accumulating pathology over lifespan (e.g., MRI, PET). Examining the relationship between olfaction and episodic memory using BSIT as a time-varying covariate (Model EP3) then determined that assessment-to-assessment variation in olfaction mirrors variation in episodic memory over time. This exciting new finding lends support to the hypothesis that BSIT could be an easier to detect proxy for AD pathology. Additionally, there is a strong positive relationship between BSIT and memory fluctuations for individuals with intermediate to high AD pathology, but not for those with low AD pathology (Figure 3). This could be due to low variability in memory, BSIT, or both in the low AD pathology individuals. The significant association of BSIT_PM level 2 predictor at baseline was surprising and could be due to the fact that person mean BSIT collapses across all occasions to provide each person's own mean score (i.e., individuals who are declining in their BSIT scores will have lower person-

mean). The significance of this term appears to support Peto's (1981) "horse-racing effect", thus further supporting not including baseline episodic memory in models where episodic memory is the outcome (Model EP0b vs EP0). However, using a time-varying-covariate (Model EP3) that changes systematically over time could be problematic (Curran & Obeidat, 2010). Since BSIT itself is changing over time, person-mean (BSIT_PM) could end up being the same for someone with average but unchanging BSIT over time and someone with initially high BSIT who declines a lot over time. A follow-up study using a bivariate model is planned in order to clarify some of these questions. To the best of our knowledge, this is the first time a growth model for episodic memory with BSIT as time-varying covariate has been explored with AD pathology and it has provided some valuable insights into the relationships between these three variables.

Population selection in MAP presents a limitation in the current study. The individuals had homogenous features (i.e., all agreed to brain donation) and therefore the generalizability of these findings may be affected. To investigate this further, these analyses could be replicated in other data sets. Head circumference was not available in this data. However, when available, it should be taken into consideration during replications as larger brain volumes have been shown to moderate cognitive decline (Erten-Lyons et al., 2013).

Intra-individual change tracking provides the most accurate information regarding an individual's health, and repeated frequent measuring from an early age would provide the most accurate information on change. A plethora of cross-sectional studies on olfaction and its relationship to cognitive decline and disease have been published, but

the nature of olfaction is that each person has a different baseline level and trajectory. Thus, longitudinal designs are required to further investigate the potential uses of olfaction as a biomarker for disease. Early indicators for disease are crucial for implementing interventions while the brain is relatively undamaged and still functioning normally.

Olfactory identification deficits have been associated with a 4–5 fold increased risk of conversion from MCI to AD (Devanand et al., 2008), and regular testing could provide a good indication of which individuals are at risk of AD, a useful index of damaged brain regions, and an effective way to track disease progression. This research strengthens the hypothesis that a decline in olfactory identification is related to AD pathology and cognitive decline. With further work in this area, olfaction could become an excellent tool for health assessments and contribute to further the knowledge of Alzheimer's disease and dementia.

References

Albers, M.W., Tabert, M.H., & Devanand, D.P. (2006). Olfactory dysfunction as a predictor of neurodegenerative disease. Current Neurology and Neuroscience Reports, 6(5), 379-386.

AlzForum, (A.R.). Alzheimer research forum: Meta-analyses of apolipoprotein e AD association studies. Retrieved March 24, 2017, from http://www.alzgene.org/meta.asp?geneID=83

Attems, J., Walker, L., & Jellinger, K.A. (2015). Olfaction and aging: A mini-review. Gerontology.

Bates, D., Maechler, M., & Bolker, B. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Bockaert, J., & Pin, J.P. (1999). Molecular tinkering of g protein-coupled receptors: An evolutionary success. The EMBO journal, 18(7), 1723-1729.

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. Acta Neuropathologica, 82(4), 239-259.

Braak, H., & Del Tredici, K. (2015). The preclinical phase of the pathological process underlying sporadic Alzheimer's disease. Brain, 138(Pt 10), 2814-2833. doi: 10.1093/brain/awv236

Brämerson, A., Johansson, L., Ek, L., Nordin, S., & Bende, M. (2004). Prevalence of olfactory dysfunction: The Skövde population-based study. The Laryngoscope, 114(4), 733-737.

Buck, L., & Axel, R. (1991). A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. Cell, 65(1), 175-187.

Campisi, J., & di Fagagna, F.d.A. (2007). Cellular senescence: When bad things happen to good cells. Nature reviews Molecular Cell Biology, 8(9), 729-740.

Croy I., Nordin, S., & Hummel, T. (2014). Olfactory disorders and quality of life—an updated review. *Chemical Senses.* 39(3):185-194.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development*, *11*(2), 121-136.

Devanand, D.P., Lee, S., Manly, J., Andrews, H., Schupf, N., Doty, R.L., . . . Mayeux, R. (2015). Olfactory deficits predict cognitive decline and Alzheimer dementia in an urban community. Neurology, 84(2), 182-189.

Devanand, D.P., Michaels-Marston, K.S., Liu, X., Pelton, G.H., Padilla, M., Marder, K., . . . Mayeux, R. (2000). Olfactory deficits in patients with mild cognitive impairment predict Alzheimer's disease at follow-up. American Journal of Psychiatry, 157(9), 1399-1405.

Djordjevic, J., Jones-Gotman, M., De Sousa, K., & Chertkow, H. (2008). Olfaction in patients with mild cognitive impairment and Alzheimer's disease. Neurobiology of Aging, 29(5), 693-706.

Doty, R.L. (2001). The Brief Smell Identification Test administration manual: Sensonics.

Doty, R.L. (2009). The olfactory system and its disorders. Paper presented at: Seminars in Neurology.

Doty, R.L., Frye, R.E., & Agrawal, U. (1989). Internal consistency reliability of the fractionated and whole University of Pennsylvania Smell Identification Test. Perception & Psychophysics, 45(5), 381-384.

Doty, R.L., Kamath V. (2014). The influences of age on olfaction: a review. *Frontiers in psychology,* 5.

Doty, R.L., Marcus, A., & William Lee, W. (1996). Development of the 12-item Cross-Cultural Smell Identification Test (CC-SIT). The Laryngoscope, 106(3), 353-356.

Doty, R.L., McKeown, D.A., Lee, W.W., & Shaman, P. (1995). A study of the test-retest reliability of ten olfactory tests. Chemical Senses, 20(6), 645-656.

Doty, R.L., Newhouse, M.G., & Azzalina, J.D. (1985). Internal consistency and short-term test-retest reliability of the University of Pennsylvania Smell Identification Test. Chem Senses, 10(3), 297-300.

Doty, R.L., Reyes, P.F., & Gregor, T. (1987). Presence of Both Odor Identification and Detection Deficits in Alzheimer's Disease. Brain Research Bulletin, 18(5), 597-600.

Doty, R.L., Shaman, P., & Dann, M. (1984). Development of the University of Pennsylvania Smell Identification Test: A standardized microencapsulated test of olfactory function. Physiol Behav, 32(3), 489-502.

Doty, R.L., Shaman, P., Kimmelman, C.P., & Dann, M.S. (1984). University of Pennsylvania Smell Identification Test: A rapid quantitative olfactory function test for the clinic. The Laryngoscope, 94(2), 176-178.

Duff, K., McCaffrey, R.J., & Solomon, G.S. (2002). The pocket smell test. The Journal of Neuropsychiatry and Clinical Neurosciences, 14(2), 197-201.

.

Eisenbach, M., & Giojalas, L.C. (2006). Sperm guidance in mammals—an unpaved road to the egg. Nature Reviews Molecular Cell Biology, 7(4), 276-285.

Elias, M.F., Beiser, A., Wolf, P.A., Au, R., White, R.F., & D'Agostino, R.B. (2000). The preclinical phase of Alzheimer disease: A 22-year prospective study of the Framingham Cohort. Archives of Neurology, 57(6), 808-813.

Eliason, S. R. (1993). *Maximum likelihood estimation: Logic and practice* (Vol. 96). Sage Publications.

Erten-Lyons, D., Dodge, H.H., Woltjer, R., Silbert, L.C., Howieson, D.B., Kramer, P., & Kaye, J.A. (2013). Neuropathologic basis of age-associated brain atrophy. JAMA Neurol, 70(5), 616-622. doi: 10.1001/jamaneurol.2013.1957

Farrer, L.A., Cupples, L.A., Haines, J.L., Hyman, B., Kukull, W.A., Mayeux, R., . . . van Duijn, C.M. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein e genotype and Alzheimer disease: A meta-analysis. Jama, 278(16), 1349-1356.

Feldmesser, E., Olender, T., Khen, M., Yanai, I., Ophir, R., & Lancet, D. (2006). Widespread ectopic expression of olfactory receptor genes. BMC genomics, 7(1), 121.

Fiandaca, M.S., Kapogiannis, D., Mapstone, M., Boxer, A., Eitan, E., Schwartz, J.B., . . . Goetzl, E.J. (2015). Identification of preclinical Alzheimer's disease by a profile of pathogenic proteins in neurally derived blood exosomes: A case-control study. Alzheimers Dement, 11(6), 600-607 e601. doi: 10.1016/j.jalz.2014.06.008

Fiandaca, M.S., Mapstone, M.E., Cheema, A.K., & Federoff, H.J. (2014). The critical need for defining preclinical biomarkers in Alzheimer's disease. Alzheimers Dement, 10(3 Suppl), S196-212. doi: 10.1016/j.jalz.2014.04.015

Filmore, D. (2004). Cell-based screening assays and structural studies are fueling g-protein coupled receptors as one of the most popular classes of investigational drug targets. Modern Drug Discovery, 7(11).

Frasnelli, J., & Hummel, T. (2005). Olfactory dysfunction and daily life. European Archives of Oto-Rhino-Laryngology and Head & Neck, 262(3), 231-235.

Frost, S., Kanagasingam, Y., Sohrabi, H., Vignarajan, J., Bourgeat, P., Salvado, O., . . . Szoeke, C. (2013). Retinal vascular biomarkers for early detection and monitoring of Alzheimer's disease. Translational psychiatry, 3(2), e233.

Goedert, M., Clavaguera, F., & Tolnay, M. (2010). The propagation of prion-like protein inclusions in neurodegenerative diseases. Trends in neurosciences, 33(7), 317-325.

Goudsmit, N., Coleman, E., Seckinger, R.A., Wolitzky, R., Stanford, A.D., Corcoran, C., . . . Malaspina, D. (2003). A Brief Smell Identification Test discriminates between deficit and non-deficit schizophrenia. Psychiatry Research, 120(2), 155-164.

Gu, X., Karp, P.H., Brody, S.L., Pierce, R.A., Welsh, M.J., Holtzman, M.J., & Ben-Shahar, Y. (2014). Chemosensory functions for pulmonary neuroendocrine cells. American Journal of Respiratory Cell and Molecular Biology, 50(3), 637-646.

Hall, C. B., Lipton, R. B., Sliwinski, M., & Stewart, W. F. (2000). A change point model for estimating the onset of cognitive decline in preclinical Alzheimer's disease. *Statistics in Medicine*, *19*(11-12), 1555-1566.

Hawkins, K.A., & Pearlson, G.D. (2011). Age and gender but not common chronic illnesses predict odor identification in older African Americans. The American Journal of Geriatric Psychiatry, 19(9), 777-782.

Hedner, M., Larsson, M., Arnold, N., Zucco, G.M., & Hummel, T. (2010). Cognitive factors in odor detection, odor discrimination, and odor identification tasks. J Clin Exp Neuropsychol, 32(10), 1062-1067. doi: 10.1080/13803391003683070

Hofer, S.M., & Piccinin, A.M. (2009). Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. Psychol Methods, 14(2), 150-164. doi: 10.1037/a0015566

Hummel, T., Sekinger, B., Wolf, S.R., Pauli, E., & Kobal, G. (1997). 'Sniffin' sticks': Olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. Chem Senses, 22(1), 39-52.

Jack, C.R., Knopman, D.S., Jagust, W.J., Shaw, L.M., Aisen, P.S., Weiner, M.W., . . . Trojanowski, J.Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. The Lancet Neurology, 9(1), 119-128.

Jack, C.R., Vemuri, P., Wiste, H.J., Weigand, S.D., Lesnick, T.G., Lowe, V., . . . Gunter, J.L. (2012). Shapes of the trajectories of 5 major biomarkers of Alzheimer disease. Archives of neurology, 69(7), 856-867.

Jack Jr, C.R., Knopman, D.S., Jagust, W.J., Petersen, R.C., Weiner, M.W., Aisen, P.S., . . . Weigand, S.D. (2013). Update on hypothetical model of Alzheimer’s disease biomarkers. Lancet Neurology, 12(2), 207.

Kim J, Basak JM, Holtzman DM. The role of apolipoprotein E in Alzheimer’s disease. Neuron. 2009;63:287–303.

Kjelvik, G., Sando, S.B., Aasly, J., Engedal, K.A., & White, L.R. (2007). Use of the Brief Smell Identification Test for olfactory deficit in a Norwegian population with Alzheimer's disease. International journal of geriatric psychiatry, 22(10), 1020-1024.

Kovács, T. (2004). Mechanisms of olfactory dysfunction in aging and neurodegenerative disorders. Ageing research reviews, 3(2), 215-232.

Krantz, E.M., Schubert, C.R., Dalton, D.S., Zhong, W., Huang, G.H., Klein, B.E.K., . . . Cruickshanks, K.J. (2009). Test–retest reliability of the San Diego Odor Identification Test and comparison with the Brief Smell Identification Test. Chemical senses, 34(5), 435-440.

Liu, C.-C., Kanekiyo, T., Xu, H., & Bu, G. (2013). Apolipoprotein e and Alzheimer disease: Risk, mechanisms and therapy. Nature Reviews Neurology, 9(2), 106-118.

Lledo P-M., Alonso M., Grubb M.S. (2006) Adult neurogenesis and functional plasticity in neuronal circuits. *Nature Reviews Neuroscience.* 7(3):179-193.

Lötsch, J., Reichmann, H., & Hummel, T. (2008). Different odor tests contribute differently to the evaluation of olfactory loss. Chemical senses, 33(1), 17-21.

Lüdecke D (2017). _sjPlot: Data Visualization for Statistics in Social Science_. R package version 2.3.0

Mackay-Sim, A., Johnston, A.N., Owen, C., & Burne, T.H. (2006). Olfactory ability in the healthy population: Reassessing presbyosmia. Chem Senses, 31(8), 763-771. doi: 10.1093/chemse/bjl019

Makowska, I., Kloszewska, I., Grabowska, A., Szatkowska, I., & Rymarczyk, K. (2011). Olfactory deficits in normal aging and Alzheimer's disease in the polish elderly population. Arch Clin Neuropsychol, 26(3), 270-279. doi: 10.1093/arclin/acr011

Mapstone, M., Dickerson, K., & Duffy, C.J. (2008). Distinct mechanisms of impairment in cognitive ageing and Alzheimer's disease. Brain, 131(Pt 6), 1618-1629. doi: 10.1093/brain/awn064

Masurkar, A.V., & Devanand, D.P. (2014). Olfactory dysfunction in the elderly: Basic circuitry and alterations with normal aging and Alzheimer's disease. Current geriatrics reports, 3(2), 91-100.

Menon, C., Westervelt, H.J., Jahn, D.R., Dressel, J.A., & O'Bryant, S.E. (2013). Normative performance on the Brief Smell Identification Test (BSIT) in a multi-ethnic bilingual cohort: A project frontier study 1. The Clinical Neuropsychologist, 27(6), 946-961.

Mesholam, R. I., Moberg, P. J., Mahr, R. N., & Doty, R. L. (1998). Olfaction in neurodegenerative disease: a meta-analysis of olfactory functioning in Alzheimer's and Parkinson's diseases. *Archives of neurology, 55*(1), 84-90.

Mobley, A.S., Rodriguez-Gil, D.J., Imamura, F., & Greer, C.A. (2014). Aging in the olfactory system. Trends in Neurosciences, 37(2), 77-84.

Murphy, C. (2002). Olfactory functional testing: Sensitivity and specificity for Alzheimer's disease. Drug Development Research, 56(2), 123-131.

Murphy, C., Schubert, C.R., Cruickshanks, K.J., Klein, B.E.K., Klein, R., & Nondahl, D.M. (2002). Prevalence of olfactory impairment in older adults. Jama, 288(18), 2307-2312.

Murphy, C., Solomon, E.S., Haase, L., Wang, M., & Morgan, C.D. (2009). Olfaction in aging and Alzheimer's disease. Annals of the New York Academy of Sciences, 1170(1), 647-657.

Nordin, S., Monsch, A. U., & Murphy, C. (1995). Unawareness of smell loss in normal aging and Alzheimer's disease: discrepancy between self-reported and diagnosed smell sensitivity. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 50(4), P187-P192.

O'Brien, R. J., & Wong, P. C. (2011). Amyloid precursor protein processing and Alzheimer's disease. *Annual review of neuroscience*, *34*, 185-204.

Ohm, T.G., & Braak, H. (1987). Olfactory bulb changes in Alzheimer's disease. Acta neuropathologica, 73(4), 365-369.

Pencina, M.J., Larson, M.G., & D'Agostino, R.B. (2007). Choice of time scale and its effect on significance of predictors in longitudinal studies. Statistics in medicine, 26(6), 1343-1359.

Peters, J.M., Hummel, T., Kratzsch, T., Lötsch, J., Skarke, C., & Frölich, L. (2014). Olfactory function in mild cognitive impairment and Alzheimer's disease: An investigation using psychophysical and electrophysiological techniques. American Journal of Psychiatry.

Peto, R. (1981). The horse-racing effect. *The Lancet*, *318*(8244), 467-468.

Pinto, J.M., Wroblewski, K.E., Kern, D.W., Schumm, L.P., & McClintock, M.K. (2014). Olfactory dysfunction predicts 5-year mortality in older adults.

Prince, M., Wimo, A., Guerchet, M., Ali, G. C., Wu, Y. T., & Prina, M. (2015). World Alzheimer Report 2015. The global impact of dementia. An analysis of prevalence, incidence, cost & trends; Alzheimer's Disease International: London.

Purves, D., Augustine, G.J., Fitzpatrick, D., Katz, L.C., LaMantia, A.-S., McNamara, J.O., & Williams, S.M. (2001). The Organization of the Olfactory System.

R Development Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA

Richens, J.L., Vere, K.-A., Light, R.A., Soria, D., Garibaldi, J., Smith, A.D., . . . Morgan, K. (2014). Practical detection of a definitive biomarker panel for Alzheimer's disease; comparisons between matched plasma and cerebrospinal fluid. International journal of molecular epidemiology and genetics, 5(2), 53.

Schubert, C.R., Cruickshanks, K.J., Fischer, M.E., Huang, G.-H., Klein, B.E.K., Klein, R., . . . Nondahl, D.M. (2011). Olfactory impairment in an adult population: The Beaver Dam Offspring Study. Chemical Senses, bjr102.

Schubert, C. R., Cruickshanks, K. J., Klein, B. E. K., Klein, R., & Nondahl, D. M. (2011). Olfactory impairment in older adults: Five-year incidence and risk factors. *The Laryngoscope, 121*(4), 873-878.

Shepherd, G.M. (2006). Smell images and the flavour system in the human brain. Nature, 444(7117), 316-321.

Shepherd, G.M. (2007). Perspectives on olfactory processing, conscious perception, and orbitofrontal cortex. Annals of the New York Academy of Sciences, 1121(1), 87-101.

Sliwinski, M. J., Hofer, S. M., Hall, C., Buschke, H., & Lipton, R. B. (2003). Modeling memory decline in older adults: the importance of preclinical dementia, attrition, and chronological age. *Psychology and aging*, *18*(4), 658.

Spehr, M., Gisselmann, G., Poplawski, A., Riffell, J.A., Wetzel, C.H., Zimmer, R.K., & Hatt, H. (2003). Identification of a testicular odorant receptor mediating human sperm chemotaxis. Science, 299(5615), 2054-2058.

Sun, G.H., Raji, C.A., MacEachern, M.P., & Burke, J.F. (2012). Olfactory identification testing as a predictor of the development of Alzheimer's dementia: A systematic review. The Laryngoscope, 122(7), 1455-1462.

Walker, L.C., Diamond, M.I., Duff, K.E., & Hyman, B.T. (2013). Mechanisms of protein seeding in neurodegenerative diseases. JAMA neurology, 70(3), 304-310.

Whitman, M. C., & Greer, C. A. (2009). Adult neurogenesis and the olfactory system. *Progress in Neurobiology*, *89*(2), 162-175.

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wilson, R.S., Arnold, S.E., Schneider, J.A., Boyle, P.A., Buchman, A.S., & Bennett, D.A. (2009). Olfactory impairment in presymptomatic Alzheimer's disease. Annals of the New York Academy of Sciences, 1170(1), 730-735.

Wilson R.S., Arnold S.E., Schneider J.A., Tang Y., Bennett D.A. (2007). The relationship between cerebral Alzheimer's disease pathology and odour identification in old age. *Journal of Neurology, Neurosurgery & Psychiatry.* 78(1):30-35.

Wilson, R.S., Arnold, S.E., Tang, Y., & Bennett, D.A. (2006). Odor identification and decline in different cognitive domains in old age. Neuroepidemiology, 26(2), 61-67.

Wilson, R.S., Schneider, J.A., Arnold, S.E., Tang, Y., Boyle, P.A., & Bennett, D.A. (2007). Olfactory identification and incidence of mild cognitive impairment in older age. Archives of general psychiatry, 64(7), 802-808.

Wilson, R.S., Yu, L., & Bennett, D.A. (2010). Odor identification and mortality in old age. Chemical Senses, bjq098.

Winberg, J., & Porter, R.H. (1998). Olfaction and human neonatal behaviour: Clinical implications. Acta Paediatrica, 87(1), 6-10.