Green Space Exposure and Cognition: The Multi-Ethnic Study of Atherosclerosis

Demitris Haldeos

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Public Health
Occupational and Environmental Medicine

University of Washington

2018

Committee:

Howard Frumkin

Robyn McClelland

Sverre Vedal

Program Authorized to Offer Degree:

School of Public Health

University of Washington

Abstract

Green Space Exposure and Cognition: The Multi-Ethnic Study of Atherosclerosis

Demitris Haldeos

Chair of Supervisory Committee:
Sverre Vedal
Environmental and Occupational Health Sciences

RATIONALE: The effects of long term green space exposure on cognition are not fully understood. With the aging of the world's population, increasing prevalence of Alzheimer's Dementia, and continued detachment from nature, we seek to contribute to quality research investigating the association of long term green space exposure with cognition.

OBJECTIVES: To determine if long term exposure to green space is associated with better cognition, measured by Normalized Difference Vegetation Index (NDVI) and Cognitive Abilities Screening Instrument, Version 2 (CASI), respectively.

METHODS: We performed analyses of participants enrolled in the MESA (Multi-Ethnic Study of Atherosclerosis) Exam 5 administered from April 2010 through February 2012. This population-based cohort was aged 45–84 years during the initial MESA study July 2000-August 2002. Time-weighted averages of NDVI in 250m and 500m buffers were calculated at individual addresses for the preceding ten years for 4,106 participants. CASI measures taken during Exam 5 were available for 4,278 participants. Regression models were used to determine whether CASI scores were associated with NDVI. Models were adjusted for age, race/ethnicity, income, education, predicted exposure to particulate matter of 2.5 micron aerodynamic diameter or less (PM2.5), apolipoprotein E allele polymorphism status and location (site). Additionally, regression models were used to determine if the presence of certain apolipoprotein E allele polymorphisms modify the association.

RESULTS: NDVI score increment of one standard deviation (SD), 33.55 points, was associated with 0.47 (95% CI .007, .021, $p<0.001$) greater CASI score, controlling for potential confounders including age, race/ethnicity, education, income, apolipoprotein E allele status, and predicted air pollution (PM2.5). With the addition of site into the model, the association of NDVI and CASI became statistically non-significant, with an effect estimate of 0.30 (95% CI -0.004, 0.022, p=.203). The apolipoprotein E4 allele modified the association between NDVI and CASI, with strongest effects found in those with two copies of the apolipoprotein E4 allele.

CONCLUSIONS: Increased proximal (250-500m) residential greenspace, as measured by NDVI, was positively associated with better performance on cognitive testing validated for dementia screening, as measured by CASI, in an elderly population (mean age 69.7 y). The association was strongest among those with two copies of the apolipoprotein E4 allele, a polymorphism associated with higher rates of developing Alzheimer's disease dementia. Limiting the NDVI comparisons to within-city contrasts resulted in attenuation of the association.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank the other investigators, the staff and the participants of the MESA study for their valuable contributions. In particular, I would like to thank Sverre Vedal, MD, MSc, Professor and Thesis Committee Chair and a wonderful advisor, Robyn McClelland, PhD, an astute Biostatistician, and Howard Frumkin, MD, DrPH, the former Dean of the University of Washington School of Public Health, for their contributions to the interpretation of the data and assistance in the editing of this manuscript. I also thank Emanuel J. Carter, and Dr. Nakatsugawa, and the Faculty of the State University of New York School of Environmental Science and Forestry, Syracuse, New York, for their inspiration and support during my completion of a master's degree in Environmental Science which served as part of the basis for my interest in this topic. I also thank the team of World Health Organization Pediatric/ Public Health Physician volunteers, and Andrew Dannenberg, MD, MPH, who encouraged me to pursue medicine. I also thank my students from the New York City Public Schools, who rekindled in me a passion for Science. A list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org.

Introduction

The population of Americans over the age of 65 is projected to increase from 40 million in 2010 to over 88 million by 2050. This aging phenomenon is global. In 2015, 8.5% of the world's population was 65 and over, a proportion that is projected to reach 17% by 2050 (He, 2015.) Normal aging is associated with decline of certain cognitive abilities, including certain memory, language and visuospatial abilities, processing speed, and executive function. Known etiology includes decrease in grey and white matter, reduced neurotransmitters, and accumulation of beta-amyloid protein which lead to neuronal death. However, with normal aging these changes are small, and should not result in functional impairment (Harada, 2013).

People who cognitively decline more than what is considered normal are diagnosed with cognitive impairment (dementia) and account for an increasingly disproportionate amount of medical costs. Alzheimer's disease (AD), the most common form of dementia, ranks sixth in cause of U.S. death, and fifth in cause of death amongst persons age 65 and older (Heron, 2009.) It is estimated that 5.7 million Americans have Alzheimer's dementia. By 2050, the number is expected to more than double to 13.8 million (Alzheimer's Association, 2018). Risk of developing AD disease doubles every five years starting at age 65. By age 85, 25% to 50% will exhibit signs of AD. In the United States, dementia costs range from \$159 billion to \$215 billion annually, costlier than either heart disease or cancer (Rand, 2013). Better evidence on what might prevent or mitigate the onset or acceleration of dementia or AD is needed.

While the etiology of dementia and AD are not very clear, it is known that the presence of the apolipoprotein E4 allele increases the risk of developing AD. In the general population, the minor allele prevalence is approximately 10-15%. Having one copy increases risk of AD by two- to threefold, while 2 copies can increase risk by ten- to twelve-fold (Liu, 2013; Michaelson, 2014). Conversely, apolipoprotein E2 is believed to impart a protective effect, while apolipoprotein E3 is thought to be neutral.

Concurrently, there is an accumulation of evidence that contact with nature can promote health (Hartig, 2014). One benefit associated with nature contact in some studies has been cognitive function. A recent study found an association between living further away from nature and lower cognitive function scores (Ziljema, 2017.) Yet, a meta-analysis of the literature (de Keijzer, 2016) reviewed over 9,000 articles and found that only 13 were of sufficient quality to investigate effects of long term green space exposure on cognition. Further, only 4 out of the 13 studies targeted older adults. It was concluded that the evidence was inadequate, in part due to the lack of objective outcome measures, as only two studies used objective tests.

With aging and continued detachment from nature (and nature's positive health impact), it is nevertheless plausible that those with less interaction with nature will have poorer cognitive outcomes. We seek to capitalize on high quality exposure and outcome data (Frumkin, 2017) from the MESA study. Previously collected NDVI (Normalized Difference Vegetation Index) data, an indicator of green space linked to residential address, objective cognitive performance scores (CASI), and individual apolipoprotein E allele status can be used to assess the association between cognitive health and exposure to green space, and whether these are modified by apolipoprotein E allele status.

Specific Aims and Hypothesis

1. Test for an association between individual-level residential exposure to green space, as measured by NDVI, and individual-level global cognitive performance, as objectively measured by CASI (Cognitive Abilities Screening Instrument Version 2). It is hypothesized that greater residential exposure to greenspace as a time-weighted average over the preceding 10 years is associated with greater cognitive performance in the elderly.

2. Determine if presence of certain apolipoprotein E alleles modify the association between individual-level exposure to green space and individual-level global cognitive performance. It is hypothesized that certain apolipoprotein E alleles will modify the association between exposure to green space and cognitive function.

Methods

Study design and participant selection

The Multi-Ethnic Atherosclerosis Study (MESA), funded by the National Heart Lung and Blood Institute to investigate subclinical cardiovascular disease, is a prospective cohort study. The MESA air pollution project (MESA Air) is an ancillary study which collected air pollution exposure estimates and NDVI based on residential addresses for the MESA cohort. Briefly, MESA enrolled participants who were free of known cardiovascular disease from six centers around the United States: Baltimore, MD; Chicago, IL; Los Angeles County, CA; New York City, NY; St Paul, MN; and Winston Salem, NC. 6,814 participants, with ages between 45-84, were recruited between 2000-2002 and completed questionnaires regarding demographics, family history, medical history, lifestyle habits and psychosocial factors (Olson, 206).

NDVI

MESA Air generated data on air pollution exposures, including predicted particulate matter of 2.5 micron mean aerodynamic diameter size and smaller (PM2.5), NDVI, and geocoded individual participant addresses. NDVI is a technique that quantifies earth surface vegetation, utilizing the measured differences between near-infrared, which vegetation reflects, and red light, which vegetation absorbs. Some of earth's man-made orbiting satellites contain sensors, such as those found on LandSat, Sentinel 2, and SPOT, capable of measuring these wavelengths. MESA reported the index on a of scale 0-255, converted from a -1 to 1 scale with higher values indicating increased verdancy. The levels of granularity chosen were 250m and 500m radial buffers, as these were the smallest available to us, and were reasonable in terms of presumed participant access to greenspace both by physical and visual means, especially for an elderly population. NDVI was obtained as a series of 16-day composite images from 2006, geocoded, linked to individual residential addresses, and a summary for 50%ile (median of entire year) was calculated from 23 averages for each location (Doop, 2015). We further extrapolated this median, by time weighting the median for each year, by time lived at each address for the ten years prior to exam 5.

CASI

During MESA Exam 5, April 2010 - February 2012, participants were administered the Cognitive Abilities Screening Instrument, Version 2, (CASI). CASI has been validated for cross-cultural applicability, screening for dementia, and monitoring for dementia progression. While studies have shown age and education to be strong predictor of CASI, this instrument has been validated to correspond with severity of dementia (Teng, 1994). CASI quantitatively assesses attention, concentration, orientation, short-term memory, long-term memory, language abilities, visual construction, list-generating fluency, abstraction, and judgment, and scores range from 0-100.

Covariates

Level of education data were obtained from MESA Exam 1; data on individual race/ethnicity, gender, income, age, site/location, BMI, and smoking status were obtained from MESA Exam 5. All participants were also genotyped for polymorphisms of the apolipoprotein E gene allele. Predicted PM2.5 at individual addresses were obtained from MESA Air in which a spatio-temporal land use regression model in a universal kriging framework was used to estimate PM2.5 concentrations (Keller, 2015). As for NDVI, time-weighted PM2.5 was calculated at individual addresses for the preceding ten years.

Statistical Analysis

All statistical tests were performed in R, version 3.3.1, using R Studio platform, version 0.99.903. A two-tailed p-value with α of 0.05 determined statistical significance.

Exposure was defined by NDVI, and CASI scores were the primary outcome. Subjects with CASI scores < 10 were excluded as these outlier CASI scores were determined to be unreasonable and likely misclassified. Despite correcting for outliers, there remained a strong left-skewed tail. This non-normality was further addressed by the use of robust standard errors in regression modeling.

Quartiles of CASI and NDVI scores were cross-classified by demographic features and other covariates. These included age, PM2.5, gender, race/ethnicity, income, education, BMI, smoking status, site, and apolipoprotein E allele polymorphisms.

Multivariable linear regression was used to examine the cross-sectional association between green space exposure (NDVI) and cognition (CASI), regressing CASI on NDVI, and effect modification by apolipoprotein E allele status. Models were adjusted for potential confounders selected on the basis of preliminary analyses (Tables 1 and 2) and included age, race/ethnicity, education, income, site location, PM2.5 and apolipoprotein E allele status. Gender, smoking status and BMI, were not included as covariates as they were not found to be associated with CASI. Seven regression models were tested, each adding additional covariates (Table 3). Additionally, we included a model that included an interaction term between the exposure (NDVI) and apolipoprotein E allele status.

RESULTS

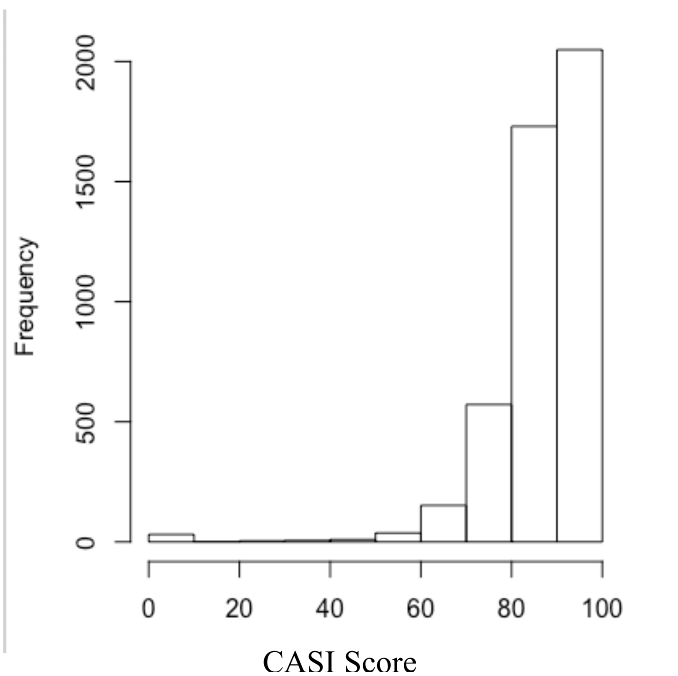
Subject Characteristics

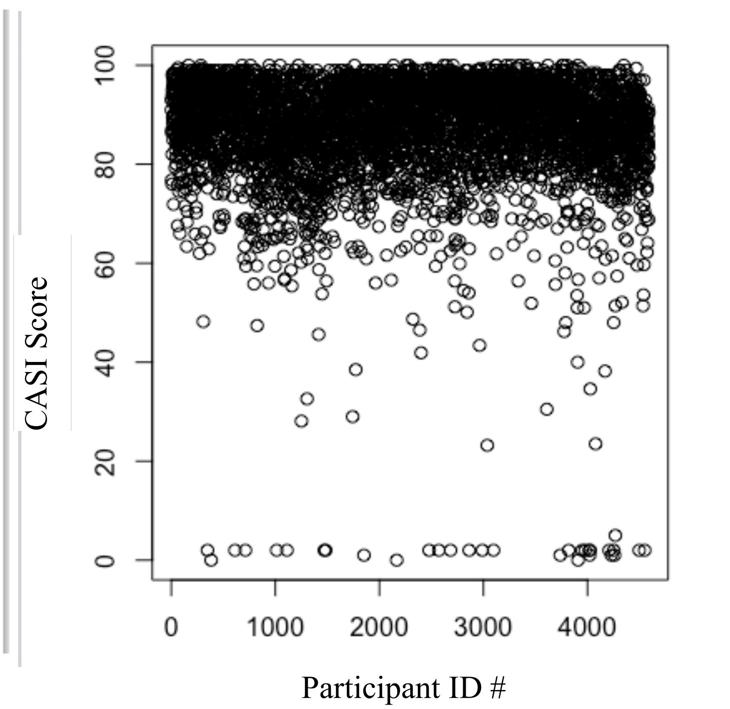
After removing 31 outliers, we analyzed 4278 MESA participants completing the CASI examination during exam 5. The sample was slightly right skewed, with a median age of 69 (IQR = 15). Informed consent was obtained for all participants and the study was approved by the Institutional Review Board at collaborating centers. Descriptive analysis was performed for CASI as below (Table 1). After combining multiple time-weighted NDVI measures and collating by identification number, we had 4106 participants. Descriptive analysis was performed for NDVI as below (Table 2).

CASI Distribution

Histogram and scatterplot of CASI scores revealed a left tailed distribution with a string of outliers below score 10, which were removed (Figures 1 and 2, respectively).

Figure 1: Histogram and Scatterplot of CASI scores, with outliers

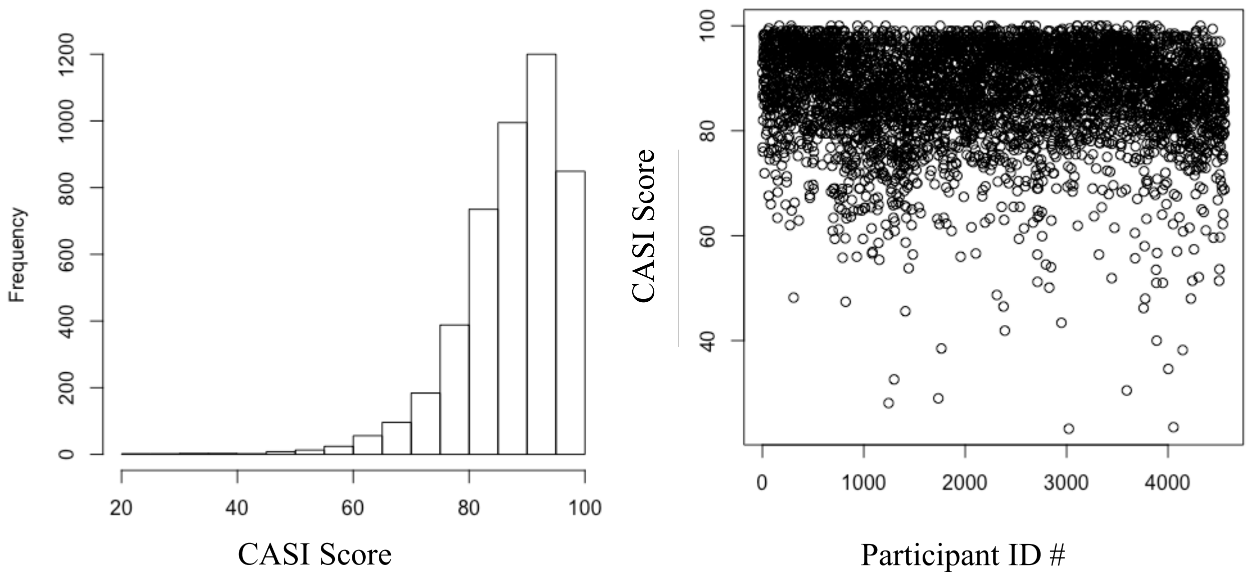




# of Participants vs. CASI score

CASI Score vs. Participant ID #

Figure 2: Histogram and Scatterplot of CASI scores, removing outliers (CASI <10)



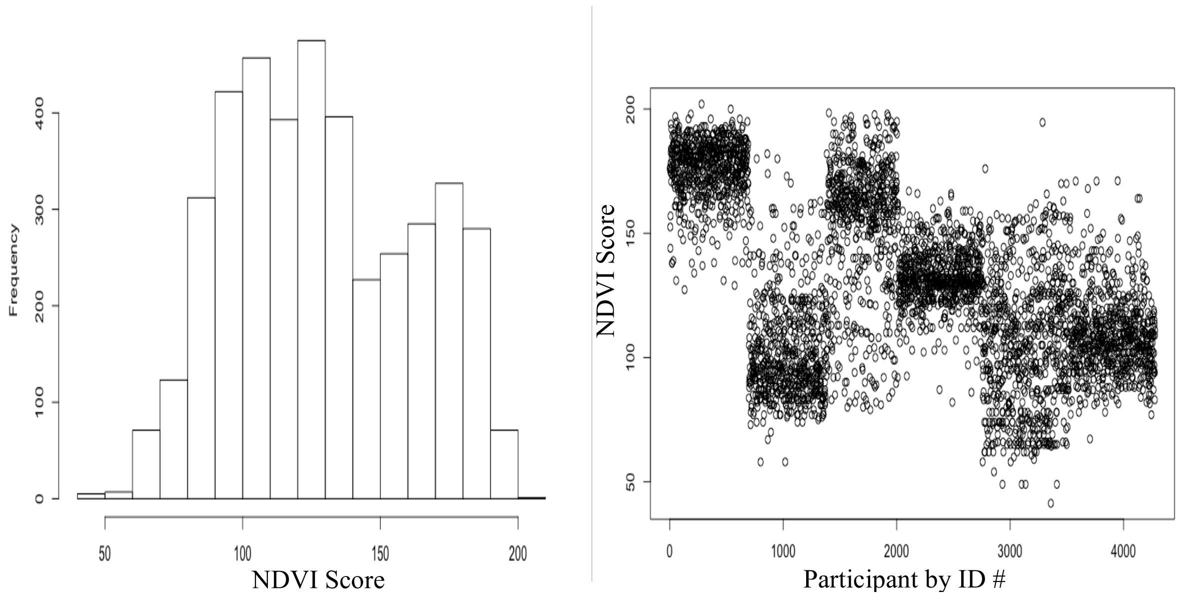
# of Participants vs. CASI score

CASI Score vs. Participant ID #

NDVI Distribution

Histogram and scatterplot of NDVI scores were constructed, for both 250m and 500m buffers. Histograms revealed a bimodal distribution, while scatterplots revealed several distinct and one amorphic agglomeration of data points for NDVI corresponding to study site. We found distributions in the 250m and 500m plots to be similar.

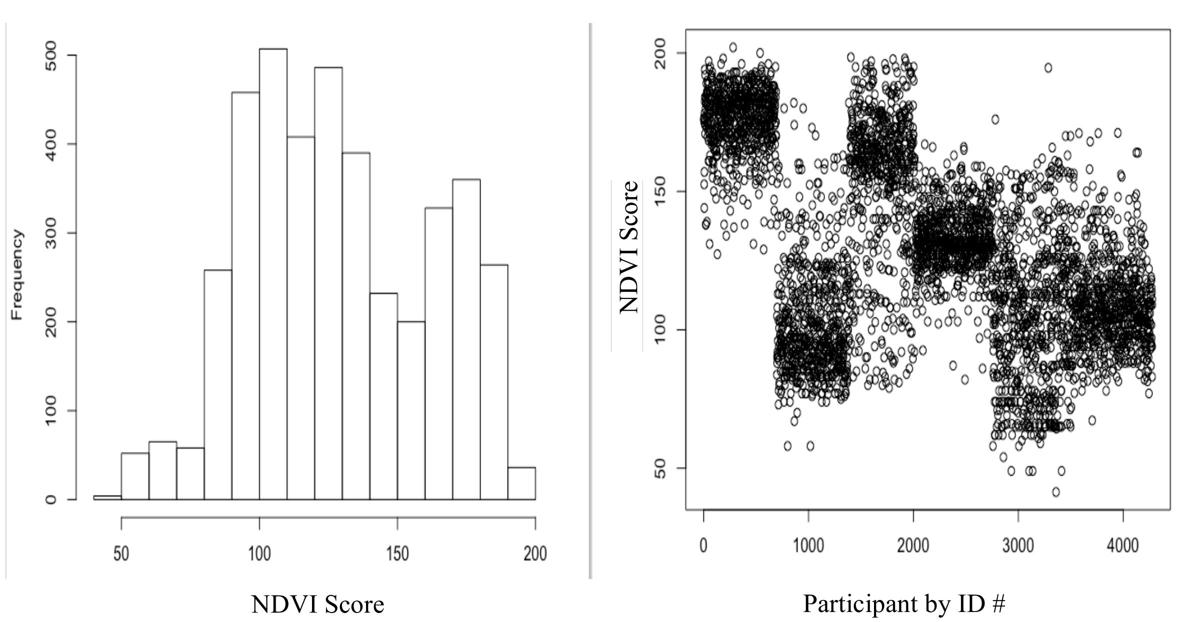
Figure 3: Histogram and Scatterplot of NDVI scores (250m buffer)



# of Participants vs. NDVI score

NDVI Score vs. Participant ID #

Figure 4: Histogram and Scatterplot of NDVI (500m)



# of Participants vs. NDVI score

NDVI Score vs. Participant ID #

<u>*CASI quartile analysis (Table 1)*</u>

CASI scores were split into quartiles, with score cutoffs at 83, 89, and 94.4. There were a little over a thousand persons in each quartile, with a total of 4278. Average age decreased with increasing CASI scores. Gender was fairly well split across quartiles of CASI scores. Race/Ethnicity was unevenly distributed across CASI quartiles, with Whites having a greater proportion in the upper quartile, and Chinese-Americans, African-Americans, and Hispanics having a greater proportion in the lowest quartile. Smoking was evenly distributed by quartile. BMI categories were evenly distributed. Apolipoprotein E allele polymorphism combinations had a mixed distribution. Those with two copies of E3 were evenly distributed while those with two copies of E2 or two copies of E4, though rarer (making up only 3% of the total data set), had CASI scores less evenly distributed. Distribution by site was uneven, with NC, MN and IL skewed towards higher quartiles of CASI. NY and CA sites are skewered towards lower quartiles of CASI. MD was evenly distributed. Distribution by income and education was markedly uneven. Lower incomes (<50,000/annum) were heavily skewed towards lower quartiles of CASI, while higher incomes (>50,000/annum) were skewed towards upper quartiles of CASI. Those with high school education or less were heavily skewed towards lower quartiles of CASI, while those with greater than high school education were skewed towards upper quartiles of CASI.

Table 1 – CASI Scores by quartile by covariates

		Q1		Q2		Q3		Q4				
		23.2-82.99		83-88.99		89-94.39		94.4-100		total		
	N=	1089	25.5%	1075	25.1%	1073	25.1%	1041	24.3%	4278	100%	
Age	Median:	74.0		70.0		68.0		65.0	median (IQR)	69 (15)	mean (sd)	70 (9)
PM2.5	Mean (SD):	13.31		13.57		13.98		14.01		13.72 (2.26)		
Gender	Female	600	26.6%	585	26.0%	538	23.9%	531	23.6%	2254	53%	
	Male	489	24.2%	490	24.2%	535	26.4%	510	25.2%	2024	47%	
Race	White	194	10.9%	373	20.9%	491	27.6%	723	40.6%	1781	42%	
	Chinese-American	170	33.9%	148	29.5%	128	25.5%	56	11.2%	502	12%	
	African-American	327	30.8%	297	28.0%	265	25.0%	173	16.3%	1062	25%	
	Hispanic	398	42.7%	257	27.5%	189	20.3%	89	9.5%	933	22%	
Smoking	Never	646	28.7%	566	25.1%	519	23.0%	523	23.2%	2254	53%	
	Quit <1Y	19	28.8%	14	21.2%	19	28.8%	14	21.2%	66	2%	
	Quit >1Y	340	21.1%	402	25.0%	438	27.2%	431	26.8%	1611	38%	
	Smoker	70	22.9%	81	26.5%	88	28.8%	67	21.9%	306	7%	
	Don't Know	11	37.9%	10	34.5%	4	13.8%	4	13.8%	29	1%	
BMI	Normal	332	26.8%	316	25.5%	285	23.0%	308	24.8%	1242	29%	
	Grade 1 (Overweight)	395	24.8%	383	24.1%	424	26.7%	388	24.4%	1591	37%	
	Grade 2 (Overweight)	319	25.1%	329	25.9%	326	25.6%	297	23.4%	1272	30%	
	Grade 3 (Overweight)	39	22.9%	47	27.6%	37	21.8%	47	27.6%	171	4%	
ApoE	e2/e2	6	17.1%	8	22.9%	14	40.0%	7	20.0%	36	1%	
	e2/e3	133	26.2%	126	24.8%	128	25.2%	121	23.8%	509	12%	
	e2/e4	30	26.1%	31	27.0%	30	26.1%	24	20.9%	116	3%	
	e3/e3	616	24.0%	653	25.5%	664	25.9%	632	24.6%	2566	61%	
	e3/e4	262	28.7%	223	24.5%	203	22.3%	224	24.6%	913	22%	
	e4/e4	32	32.7%	20	20.4%	26	26.5%	20	20.4%	99	2%	
Study Site	Winston Salem, NC	122	17.6%	154	22.2%	184	26.5%	234	33.7%	695	16%	
	New York, NY	273	39.4%	196	28.3%	121	17.5%	103	14.9%	694	16%	
	Baltimore, MD	142	22.8%	163	26.1%	159	25.5%	160	25.6%	625	15%	
	St. Paul, MN	169	22.7%	161	21.7%	197	26.5%	216	29.1%	744	17%	
	Chicago, IL	143	18.5%	167	21.7%	224	29.1%	237	30.7%	772	18%	
	Los Angeles, CA	240	31.9%	234	31.1%	188	25.0%	91	12.1%	754	18%	
Income	0-24,999	499	43.5%	306	26.7%	228	19.9%	113	9.9%	1151	28%	
	25-49,999	294	26.3%	310	27.8%	287	25.7%	225	20.2%	1119	27%	
	50-99,999	174	15.3%	272	23.9%	329	28.9%	365	32.0%	1141	27%	
	100,000+	66	8.8%	158	21.1%	203	27.1%	322	43.0%	751	18%	
Education	No Schooling	23	88.5%	2	7.7%	1	3.8%	0	0.0%	27	1%	
	Grade 1-8	228	69.1%	70	21.2%	27	8.2%	5	1.5%	331	8%	
	Grade 9-11	149	62.3%	54	22.6%	32	13.4%	4	1.7%	240	6%	
	High School/GED	233	31.3%	240	32.2%	181	24.3%	91	12.2%	746	17%	
	Some College	139	19.7%	184	26.1%	201	28.5%	182	25.8%	707	17%	
	Technical School Grad	80	25.2%	80	25.2%	100	31.4%	58	18.2%	319	7%	
	Associate Degree	47	20.3%	63	27.3%	68	29.4%	53	22.9%	232	5%	
	Bachelor's Degree	108	13.6%	187	23.5%	233	29.3%	268	33.7%	797	19%	
	Graduate/Prof Degree	80	9.1%	193	21.9%	228	25.9%	379	43.1%	881	21%	

NDVI quartile analysis (Table 2)

NDVI was split into quartiles, with score cutoffs at 102, 126, and 158. Initial analysis revealed little difference between 250m and 500m buffers, so we continued with and report this analysis for 250m only. There were a little over one thousand persons in each quartile, with a total of 4106. Average age was lower for higher NDVI scores. Gender was fairly evenly distributed across quartiles of NDVI scores. Race/Ethnicity was unevenly distributed across NDVI quartiles, with Whites and African-Americans occupying a greater proportion of the upper quartile, whilst Chinese-Americans and Hispanics had a greater proportion in the lowest quartile. Smoking was evenly distributed by quartile. BMI categories were evenly distributed. Apolipoprotein E allele polymorphism combinations had a mixed distribution. Distribution by site was markedly uneven as also shown in the scatterplots above in figures 3 and 4. Distribution by income and education was also markedly uneven, with those with higher income and education levels residing in areas with higher NDVI.

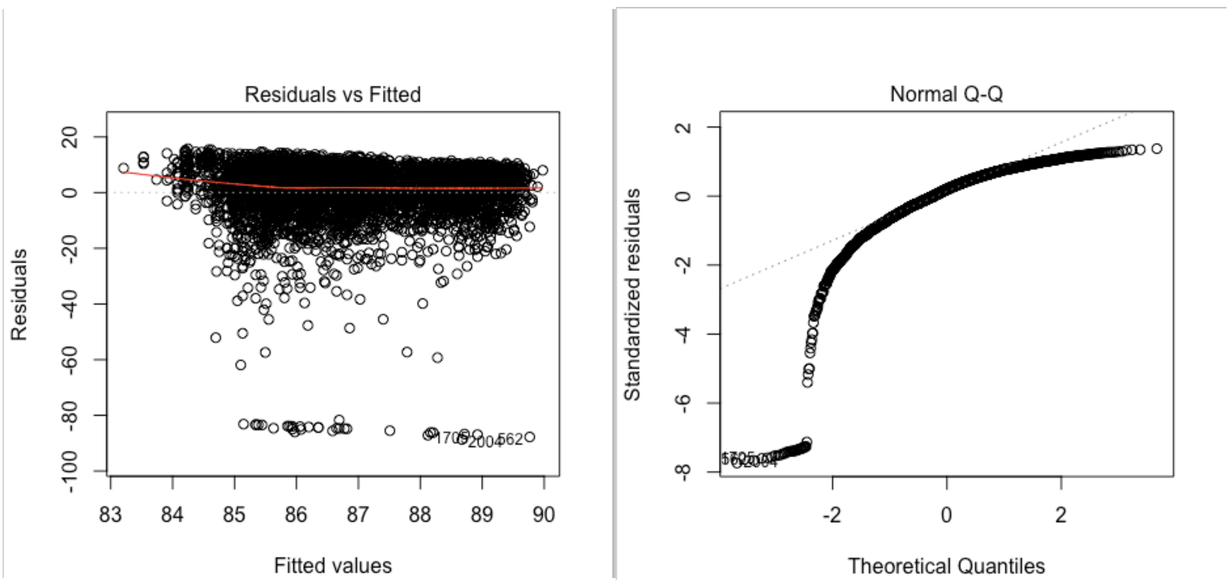
Table 2 – NDVI quartiles by covariates

Descriptive Statistics - NDVI by quartiles		Q1		Q2		Q3		Q4			
		41.4-102		102-126		126-158		158-202		total	
	N=	1027	25.0%	1035	25.2%	1028	25.0%	1016	24.7%	4106	100%
PM2.5	Mean (SD):	11.84		13.03		14.49		15.46		13.70 (1.59)	
Gender	Female	564	26.1%	550	25.4%	523	24.2%	528	24.4%	2165	53%
	Male	463	23.9%	485	25.0%	505	26.0%	488	25.1%	1941	47%
Race	White	353	20.5%	301	17.4%	498	28.9%	574	33.3%	1726	42%
	Chinese-American	166	34.7%	186	38.9%	120	25.1%	6	1.3%	478	12%
	African-American	194	19.0%	235	23.0%	166	16.3%	425	41.7%	1020	25%
	Hispanic	314	35.6%	313	35.5%	244	27.7%	11	1.2%	882	21%
Smoking	Never	579	26.7%	573	26.4%	526	24.3%	490	22.6%	2168	53%
	Quit <1Y	18	27.7%	12	18.5%	18	27.7%	17	26.2%	65	2%
	Quit >1Y	364	23.5%	374	24.2%	399	25.8%	409	26.5%	1546	38%
	Smoker	59	20.6%	69	24.0%	70	24.4%	89	31.0%	287	7%
	Don't Know	6	21.4%	6	21.4%	7	25.0%	9	32.1%	28	1%
BMI	Normal	347	29.1%	331	27.8%	261	21.9%	253	21.2%	1192	29%
	Grade 1 (Overweight)	377	24.8%	369	24.2%	396	26.0%	380	25.0%	1522	37%
	Grade 2 (Overweight)	252	20.7%	299	24.5%	334	27.4%	335	27.5%	1220	30%
	Grade 3 (Overweight)	49	29.5%	36	21.7%	37	22.3%	44	26.5%	166	4%
ApoE	e2/e2	5	14.7%	9	26.5%	13	38.2%	7	20.6%	34	1%
	e2/e3	123	25.3%	110	22.6%	132	27.1%	122	25.1%	487	12%
	e2/e4	24	21.6%	30	27.0%	21	18.9%	36	32.4%	111	3%
	e3/e3	609	24.7%	621	25.2%	638	25.9%	593	24.1%	2461	60%
	e3/e4	262	28.7%	223	24.5%	203	22.3%	224	24.6%	912	22%
	e4/e4	27	27.6%	27	27.6%	13	13.3%	31	31.6%	98	2%
Study Site	Winston Salem, NC	0	0.0%	1	0.1%	54	8.0%	620	91.9%	675	16%
	New York, NY	390	59.5%	188	28.7%	61	9.3%	17	2.6%	656	16%
	Baltimore, MD	43	7.1%	60	9.9%	155	25.5%	350	57.6%	608	15%
	St. Paul, MN	7	1.0%	166	23.2%	531	74.3%	11	1.5%	715	17%
	Chicago, IL	345	46.6%	238	32.1%	146	19.7%	12	1.6%	741	18%
	Los Angeles, CA	242	34.0%	382	53.7%	81	11.4%	6	0.8%	711	17%
	0-24,999	379	34.8%	334	30.7%	218	20.0%	157	14.4%	1089	27%
	25-29.99	238	22.1%	272	25.3%	300	27.9%	266	24.7%	1076	27%
	50-74.99	203	18.4%	257	23.3%	303	27.5%	340	30.8%	1103	28%
	100-124.99	187	26.2%	155	21.7%	171	23.9%	202	28.3%	715	18%
Education	No Schooling	11	42.3%	12	46.2%	3	11.5%	0	0.0%	27	1%
	Grade 1-8	144	45.1%	110	34.5%	55	17.2%	10	3.1%	320	8%
	Grade 9-11	73	31.9%	68	29.7%	53	23.1%	35	15.3%	230	6%
	High School/GED	158	22.1%	172	24.1%	202	28.3%	182	25.5%	715	17%
	Some College	149	21.9%	162	23.8%	178	26.2%	191	28.1%	681	17%
	Technical School Grad	49	15.8%	91	29.4%	90	29.0%	80	25.8%	311	8%
	Associate Degree	47	21.8%	47	21.8%	66	30.6%	56	25.9%	217	5%
	Bachelor's Degree	181	23.8%	166	21.8%	184	24.2%	229	30.1%	761	19%
	Graduate/Prof Degree	215	25.4%	206	24.4%	196	23.2%	228	27.0%	846	21%

Regression Analysis

Initial regression model diagnostics included plots of residuals vs fitted values and q-q plots. Evaluating the crude model prior to any data cleaning revealed non-normality, consistent with the known heavy left skew, and outliers.

Figure 5: Residual and Q-Q plot: crude model (prior to data cleaning)



After cleaning CASI data by removing 31 outliers (score <10), residual and q-q plots were repeated, with improvement in the q-q plot, but remaining heteroscedascity. Further improvement was seen with our fully adjusted model (see figures 6 and 7). Nevertheless, because of persistent skewness of the model residuals and heteroscedasticity, we elected to report regression results using robust standard errors.

Figure 6: Residual and Q-Q plot: crude model (removing outlier scores <10)

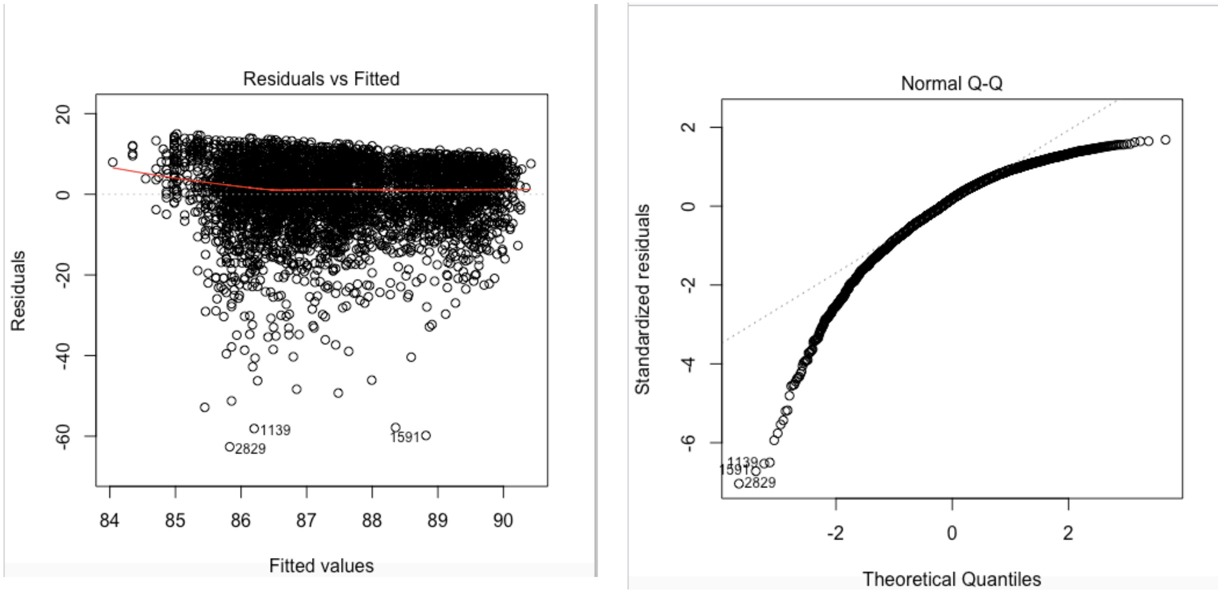
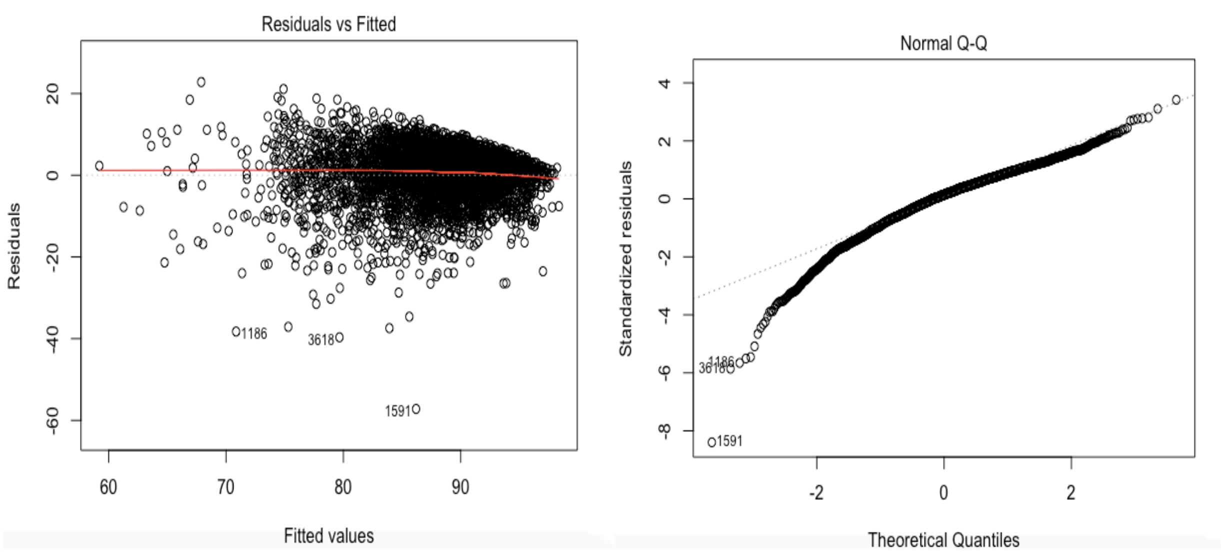


Figure 7: Residual and Q-Q plot: fully adjusted model



Simple and multivariable analysis

In a simple linear regression of untransformed CASI scores on NDVI (250m buffer), for each one standard deviation (SD) increment of NDVI there is an associated CASI score increase of 1.33 (95%CI, 1.067, 1.597, $p<.001$). For a 500m NDVI buffer, for each one SD increment of NDVI there is an associated CASI score increase of 1.22 (95%CI, 0.956, 1.487, $p<.001$). See Table 3. We determined that estimate differences between models utilizing 250m and 500m buffers were too small to be significant, and continued remaining analysis using the 250m radius buffer.

To the crude analysis, potential confounders were added one by one, noting changes in effect estimates, with 95%ile confidence intervals utilizing robust standard errors. See Table 3.

In a multiple linear regression, untransformed CASI scores were regressed on NDVI (250m buffer), after controlling for age, race/ethnicity, education, income, Apolipoprotein E allele status, and predicted PM2.5 levels. For each SD increment in NDVI there is an associated CASI score increase of .47 (95%CI, 0.235, 0.705, $p<.001$). See Table 3. When adding site location as a covariate, the effect estimate was even further attenuated and the standard error increased, resulting in a wider confidence interval and loss of statistical significance.

Table 3: Regression models with robust standard errors

	Models (+ adjustments)	NDVI*	95%CI (Low)	95%CI (High)	SE**	pval
1	Crude	1.33	1.067	1.597	0.134	<.001
2	Crude+age	1.29	1.040	1.540	0.127	<.001
3	Model 2 + race	0.51	0.275	0.735	0.117	<.001
4	Model 3 + education	0.38	0.164	0.594	0.107	=.001
5	Model 4 + income	0.33	0.114	0.540	0.107	=.006
6	Model 5 + ApoE	0.33	0.114	0.540	0.107	=.006
7	Model 6 +PM2.5	0.47	0.235	0.705	0.121	<.001
8	Model 7 + Site	0.30	-0.134	0.738	0.228	=.203
	*NDVI = Effect Estimate x 33.55 (NDVI SD)			**SE = Robust SE x 33.55 (NDVI SD)		

ApoE interaction analysis

To formally test for an interaction, ApoE allele by NDVI interaction terms were added to a model of CASI score on NDVI (250m buffer) controlling for age, race/ethnicity, education, income, apolipoprotein E allele status, and PM2.5 level. Cross-product terms reflecting both the dominant (any e4 allele) and recessive (homozygous e4/e4) models were added separately to the fully-adjusted model. Neither interaction term was statistically significant (term with any e4 allele, $p=0.48$; term with two copies [homozygous] of the e4 allele, $p=0.12$). We stratified the regression analysis for the fully adjusted model by the number of e4 alleles present (0, 1, or 2 copies). The positive association with NDVI found earlier was amplified in those participants with one or more e4 alleles. The association with NDVI in those with two copies of the e4 allele was not statistically significant due to the small number of participants in that stratum. See Table 4.

Table 4: Regression models stratified by ApoE4 (NDVI beta * SD)

copies E4	beta (adj)	95%CI lower	95%CI upper	pval	n
0	0.34	0.10	0.67	0.01	3108
1	0.67	0.34	1.34	0.01	1027
2	1.34	1.01	3.36	0.28	98
					4233

DISCUSSION

In an elderly population with mean age of 69 years, having a higher residential time-weighted NDVI score was associated with a higher CASI score. CASI scores are known to be useful in screening for, and are correlated with, severity of dementia. This positive association was maintained while controlling for age, race/ethnicity, income, education, Apolipoprotein E allele status, and predicted PM2.5 exposure. The estimate of the effect size was sensitive to control for site, attenuating the estimate and resulting in loss of statistical significance. Effect modification by genotype was found, with strengthening of the estimate in those with copies of the apolipoprotein E4 allele. Each additional copy of the E4 allele is associated with additional increase in size of the effect estimate.

This study is novel in that it tests for effect modification by the apolipoprotein E allele of the association between green space exposure and cognition. Further, we know of no other study that also controls for air pollution as specified by predicted PM2.5. This study contributes to the literature in a way called for in a recent meta-analysis (deKeijzer, 2016) by using objective outcome measures like CASI.

CASI scores are a standardized test of cognition, validated in their use with dementia, and have been shown to correlate well with severity of dementia disease. They have been used in screening for dementia disease and have been tested across ethnicities (Teng, 1994). We found CASI to be most associated with age, race/ethnicity, income, education and site.

NDVI exposure measures are also objective. Time-weighted averages of NDVI in 250m and 500m buffers were calculated at individual addresses for the ten years prior to the Exam 5 CASI test, allowing us to employ a long-term estimate of exposure. NDVI scores were most associated with age, race/ethnicity, income, education, and site. In the analysis, we controlled for variables associated with both CASI and NDVI.

The significance of our finding is particularly interesting given the ten- to twelve-fold increased risk for developing Alzheimer's dementia in those with two copies of the ApoE E4 allele. The suggestion that residential greenness may be associated with somewhat better cognitive function scores, especially in those at highest genetic risk for the development of or progression of dementia. This is worthy of further investigation, especially given the large disease burden of Alzheimer's dementia. From a preventive standpoint, geriatric medicine might utilize this knowledge in a prescriptive way, possibly encouraging patients at risk to increase their exposures to greenspace. This finding might also be utilized in a population centered way, driving long term city and urban planning and development in a direction that seeks to attenuate medical costs to society by emphasizing greenspace.

Limitations:

While NDVI is objective, it is still a crude measure of greenspace. Despite using time-weighted averages of NDVI at individual addresses for the ten years prior, there is most certainly some error in the measurement of exposure. Also, while we chose the smallest radial buffer available to us, this assumed that participants interacted with the measured level of greenness. We also acknowledge that we did not capture whether participants routinely travelled outside this area in their day to day lives. Further, participants may be able to view greenspace but not travel within it, or may travel within it but have no view from their home. We also do not have any measure of what persons did within greenspace. The predicted PM2.5 used in this study was only an estimate of outdoor residential concentration. To the extent that time spent indoors and in other micro-environments changes exposure to PM2.5, the predicted PM2.5 used here is also mismeasured.

Further, only one set of NDVI measures (from 2006) was available. Time-weighting simply incorporated moves of residence during the ten-year period. One of the limitations of a cross-sectional study is the inability to confirm the temporality of the observed associations. That is, does living in higher NDVI addresses lead to higher CASI scores, or do those with higher CASI scores tend to seek out and live in closer proximity to greenspace? Also, the appropriate time window of exposure that is relevant for affecting cognitive decline is not known. If green space exposure outside the time frame used in this study is more relevant, additional exposure measurement error could have been introduced. Finally, differential exposure measurement error could conceivably have been present if, for example, the elderly with the lowest CASI scores spent more time indoors and therefore had little green space exposure, regardless of their NDVI measure.

Control for site location resulted in attenuation of the effect of NDVI on CASI and substantial loss of precision of the effect estimate. As seen in the scatterplots of NDVI, NDVI is strongly associated with city; the association between site and CASI was not nearly as strong. Controlling for site resulted in the analysis being essentially a within-city analysis, with substantial loss of the variability in NDVI that was present between cities. This reduction in exposure variability likely explains the impact of control for city on the observed associations. On the other hand, if one is not confident that differences between city that could potentially confound the association between NDVI and CASI have not been adequately captured by the set of variables included in the fully adjusted model, then control for city would be necessary to control for such unmeasured variables. This uncertainty introduces another limitation to the interpretation of the study findings.

Regarding the interaction by ApoE genotype, because of the small number of subjects who were homozygous for the e4 allele, we had limited power to assess the interaction. Studies in other populations with a higher prevalence of the e4 allele might better address the effect of ApoE polymorphisms.

In spite of these limitations, there are great strengths in this study. We had access to a robust population dataset. The size of the population allowed for analysis with significant power. Both exposure and outcome measures are objective. The outcome measure has been validated for use in dementia screening and correlated with disease severity. We were also able to introduce interesting covariates, including apolipoprotein E allele polymorphisms and predicted PM 2.5, and uncovered an important interaction between the apolipoprotein E allele polymorphisms and the association between greenspace and cognition. The finding that residing in greenspace may be associated with improved cognition on a screening test used to screen for and assess the severity of Alzheimer's dementia, supports further investigation into the utilization of nature in a preventive fashion. Additionally, the finding that this association may be stronger in those with the highest genetic predisposition to Alzheimer's dementia, further supports this investigation, as the predicted burden of Alzheimer's dementia is so great.

CONCLUSION

Increased proximal (250-500m) residential greenspace, as measured by NDVI, was positively associated in an elderly population with better performance on cognitive testing validated for dementia screening as measured by CASI. The association was strongest among those with two copies the apolipoprotein E4 allele, a polymorphism associated with higher rates of developing Alzheimer's disease dementia.

REFERENCES

Alzheimer's Association Report (2018), Alzheimer's and Dementia, Volume 14, Issue 3, March 2018, 367-429, https://doi.org/10.1016/j.jalz.2018.02.001

DOOP (2015), Data Organization and Operating Procedures, MESA Study, http://deohs.washington.edu/sites/default/files/images/MESAAirDOOPRevised.pdf

Dadvand, P., e. (2015), Green spaces and cognitive development in primary schoolchildren. Proc. Natl. Acad. Sci. 112, 201503402. http://dx.doi.org/10.1073/pnas.1503402112.

de Keijzer, Carmen, e. (2016) Long-Term Green Space Exposure and Cognition Across the Life Course: A Systematic Review, Curr Envir Health Rpt, 3:468–477.

Diez-Roux, Ana V., e., (2016), The Impact of Neighborhoods on CV Risk, Global Heart, Vol.11, No.3, September 2016, 353-363.

Farrer, Lindsay, e. (1997), Effects of Age, Sex, and Ethnicity on the Association Between Apolipoprotein E Genotype and Alzheimer Disease. A Meta-analysis, JAMA. 1997;278(16):1349-1356.

Fitzpatrick, Anne, e. (2015), Sociodemographic Correlates of Cognition in the Multi-Ethnic Study of Atherosclerosis (MESA), Am J Geriatr Psychiatry. 2015 July ; 23(7): 684–697.

Frumkin, Howard, e. (2017), Nature Contact and Human Health: A Research Agenda, Environ Health Perspect; 2017 July: 125(7).

Harada, Caroline, e. (2013), Normal Cognitive Aging, Clin Geriatr Med. 2013 November: 29(4): 737–752.

Hartig, T., e. (2014), Nature and Health. Annu. Rev. Public Health. 35, 207–228. http://dx.doi.org/10.1146/annurev-publhealth-032013-182443.

He, Wan, e. (2015), U.S. Census Bureau, International Population Reports, P95/16-1, An Aging World: 2015, U.S. Government Publishing Office, Washington, DC, 2016.

Hendrie, Hugh, e. (2014), APOE ε4 and the risk for Alzheimer disease and cognitive decline in African Americans and Yoruba, Int Psychogeriatr. 2014 June ; 26(6): 977–985.

Hirsch, Jana, e., (2013), Discrete land uses and transportation walking in two U.S. cities: The Multi-Ethnic Study of Atherosclerosis, Health &Place: 24(2013), 196–202.

Keller JP, Olives C, Kim S-Y, Sheppard L, Sampson PD, Szpiro AA, Oron AP, Vedal S, Kaufman JD. (2015) A unified spatiotemporal modeling approach for prediction of multiple air pollutants in the Multi-Ethnic Study of Atherosclerosis and Air Pollution. Environ Health Perspect 2015; 123: 301-309. DOI: 10.1289/ehp.1408145

Liesi, E., e., (2013), Alzheimer disease in the United States (2010–2050) estimated using the 2010 census, Neurology. 2013 May 7; 80(19): 1778–1783.
doi: 10.1212/WNL.0b013e31828726f5

Liu, C., e. (2013), Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol 9(2):106-18.

Michaelson, DM, (2014), APOE ε4: the most prevalent yet understudied risk factor for Alzheimer's disease., Alzheimer's Dement 10(6):861-8.

Olson, Jean, e., (2016) Legacy of MESA: Introduction and Perspective. Glob Heart. 2016 Sep. 11(3): 269–274.

Paddick, S., e., (2017), Cognitive screening tools for identification of dementia in illiterate and low-educated older adults, a systematic review and meta-analysis. Int Psychogeriatr. 2017 Jun;29(6): 897-929. doi: 10.1017/S1041610216001976.

Rand, 2013. https://www.rand.org/pubs/external_publications/EP50247.html

Teng, EL, (1994), The Cognitive Abilities Screening Instrument (CASI): a practical test for cross-cultural epidemiological studies of dementia, Int Psychogeriatr. 1994 Spring;6(1):45-58; discussion 62

UN, United Nations Department of Economic and Social Affairs, (2015)
http://www.un.org/en/development/desa/news/population/2015-report.html

UN, World Population Ageing, (2015),
http://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf

UN, World Urbanization Prospects, (2014),
http://www.un.org/en/development/desa/publications/2014-revision-world-urbanization-prospects.html

World Health Organization (WHO), 2016, Urban green spaces and health. Copenhagen: WHO Regional Office for Europe, 2016.
http://www.euro.who.int/__data/assets/pdf_file/0005/321971/Urban-green-spaces-and-health-review-evidence.pdf?ua=1

Ziljema, Wilma, e., (2017), The relationship between natural outdoor environments and cognitive functioning and its mediators, Environmental Research 155 (2017) 268-275.