Supplementary Online Content

Deshpande SK, Hasegawa RB, Rabinowitz AR, et al. Association of playing high school football with cognition and mental health later in life. *JAMA Neurol.* Published online July 3, 2017. doi:10.1001/jamaneurol.2017.1317

eAppendix. Supplemental Information
eFigure. Comparison of Post-Matching Distributions of Important Baseline Covariates and Post-Exposure Variables by Treatment and Control Group
eTable 1. Composition of Matched Sets
eTable 2. Sports Played by Eligible Participants
eTable 3. Availability of Letter Fluency, Delayed Word Recall, and Modified CES-D Scores at Age 65 Years
eTable 4. Mean Baseline Covariates Before and After Matching Football Players to Non-Sport Controls (Analogous to Table 1 in Main Text)
eTable 5. Mean Baseline Covariates Before and After Matching Football Players to Non-Collision Sport Controls (Analogous to Table 1 in Main Text)
eTable 6. Mean Baseline Covariates Before and After Matching Non-Collision Sport Controls to Non-Sport Controls (Analogous to Table 1 in Main Text)
eTable 7. Percentages of Missing Baseline Data Before and After Matching Football Players to All Controls
eTable 8. Percentages of Missing Baseline Data Before and After Matching Football Players to Non-Sport Controls
eTable 9. Percentages of Missing Baseline Data Before and After Matching Football Players to Non-Collision Sport Controls
eTable 10. Percentages of Missing Baseline Data Before and After Matching Non-Collision Sport Controls to Non-Sport Controls
eTable 11. Analog to eTable 1, Restricted to Participants With Available *ApoE* Genotype Data
eTable 12. Analog of eTable 3, Restricted to Participants With Available *ApoE* Genotype Data
eTable 13. Analog of Table 1 in Main Text, Restricted to Participants With Available *ApoE* Genotype Data
eTable 14. Analog of eTable 4, Restricted to Participants With Available *ApoE* Genotype Data
eTable 15. Analog of eTable 5, Restricted to Participants With Available *ApoE* Genotype Data
eTable 16. Analog of eTable 6, Restricted to Participants With Available *ApoE* Genotype Data
eTable 17. Analog to eTable 7, Restricted to Participants With Available *ApoE* Genotype Data
eTable 18. Analog to eTable 8, Restricted to Participants With Available *ApoE* Genotype Data
eTable 19. Analog to eTable 9, Restricted to Participants With Available *ApoE* Genotype Data
eTable 20. Analog to eTable 10, Restricted to Participants With Available *ApoE* Genotype Data
eTable 21. Analog to Table 2, Restricted to Participants With Available *ApoE* Genotype Data
eTable 22. Analog to Table 3, Restricted to Participants With Available *ApoE* Genotype Data
eTable 23. Estimated Interaction Effect (CI) of Playing High-School Football and *ApoE* e4 Status
eTable 24. Adjusted Differences in Average Cognitive Test and CES-D Scores Over Time for Football Players and Controls
eReferences

This supplementary material has been provided by the authors to give readers additional information about their work.

© 2017 American Medical Association. All rights reserved.

eAppendix. Supplemental Information

Eligibility and Inclusion Criterion and Available Outcomes

The WLS dataset consists of 10,317 subjects, 4,991 (48.38%) of whom were males. We removed all female subjects from our analysis. To determine whether a male subject played high school football, we use data recorded from their senior year high school yearbook. The yearbook review took place between 2009 and 2011 and was performed by the WLS. In all, we are missing yearbook information for 621 males (12.44% of all male subjects) and an additional 397 (7.95%) males had yearbook information available but did not have activity participation recorded under their senior photo or in an index[1]. Since it is possible to suffer repetitive head trauma in sports such as soccer, hockey, lacrosse, and wrestling, we excluded those subjects who played one of these "collision" sports. eTable 1 shows the breakdown of sports participation among the 3,973 subjects. In all, we end up excluding another 69 (1.38%) subjects who played one of these collision sports. This left us with a total of 3,904 subjects (78.22% of all male subjects).

Of the 3,904 subjects eligible for the analysis, 1,153 (29.53%) played football (our exposed group) and 2,751 who did not play football or any other collision sport (our "all controls" group). 1,951 control subjects did not play any high school sport (our "non-sport control" group) and the remaining 800 controls played a non-contact sport (our "non-collision sport" group).

Measures of Cognition and Depression Collected by the WLS:

Between 2003 and 2005, when subjects were about 65 years old, the WLS administered six cognitive tests by telephone, five more than they administered in 1993 when subjects were in their mid-fifties. The test battery included measures of abstract and quantitative reasoning, memory/attention, and verbal fluency. The six tests were: Immediate and Delayed Word Recall (IWR and DWR), Digit Ordering (DO), WAIS-R Similarities (SIM), and Letter and Category Fluency (LF and CF). See Yonker (2007) for further details. There is broad consensus that the most commonly effected neurocognitive domains following traumatic brain injury (TBI) are attentional deficits, speeded processing, verbal recall[1]. Untimed tasks that rely on verbal and quantitative abstract reasoning skills are not typically affected by TBI, hence we determined that the SIM task was not appropriate for addressing the hypotheses of the present study. This left three measures of memory/attention and two measures of verbal fluency to be considered as outcome measures.

We selected one measure from each of these domains for our primary outcome—delayed word recall as a measure of memory/attention, and letter fluency as a measure of verbal fluency. Letter fluency was chosen among the verbal fluency tasks because it is more difficult than category fluency tasks, and performance on this task is thought to rely more on speeded processing and executive function (cognitive domains often impaired by TBI) and less on semantic knowledge (which is typically unaffected by TBI) than category fluency[1]. Furthermore, a meta-analysis of the neuropsychological effects of mTBI identified letter fluency as the task most sensitive to mTBI[2].

Delayed word recall was chosen among the tests of memory attention, in part, because this test is analogous to the Rey Auditory Verbal Learning Test- Delayed Recall, which is considered a core TBI outcome measure recommended as a Common Data Elements (CDE) in Traumatic Brain Injury research[3]. We considered this task superior to the Digit Ordering test, which is analogous to the WAIS Digit Span (DS) task for several reasons. Firstly, verbal memory tests are considered core outcome measures for studying neurocognitive sequelae of TBI, whereas DS is classified as a supplementary measure[3]. A meta-analysis of the neurocognitive effects of sports-related concussion suggests that delayed memory measures are more sensitive to concussion than measures of attention only[4]. This is possibly because performance on delayed memory tasks relies on both attention and verbal recall, and hence, assesses two cognitive skills often affected by TBI. Furthermore, impairments in delayed verbal recall may

[1] These students came from so-called "complex schools" for which activity information was not listed under senior photos or as part of an index. Instead, WLS coders had to rely on group pictures of teams and clubs to impute participation data. We find that there is no significant difference in the size of complex and non-complex schools but that the rate of football playing in complex schools is half that in non-complex schools. We suspect that there is misclassification of football playing in complex schools and we therefore drop students from these schools from our analysis.

© 2017 American Medical Association. All rights reserved.

be particularly relevant to later-life cognitive deficits associated with chronic head injury effects, as prior research suggests associations between sports-related concussion and pre-clinical Alzheimer's Disease, also known as Mild Cognitive Impairment (MCI)[5].

For our primary analysis, we formed a composite cognition score by averaging the z-scores from the LF and DWR tests. These z-scores were computed using data from all eligible subjects. While our primary analysis focused on a composite cognition score, we looked at each test individually in our secondary analyses, with the exception of CF. This is because the CF test was administered to only a subset of WLS subjects interviewed in the 2003-05 wave.

The CES-D is a 20-question self-report measure of depressive experiences[6]. Sixteen of the questions are negatively worded (e.g. "On how many days during the past week did you think your life had been a failure?") and four questions are positively worded (e.g. "On how many days during the past week did you feel happy?"). The standard CES-D scoring method first collapses responses to these questions into < 1 day, 1 – 2 days, 3 – 4 days, and 5 – 7 days, which are scored 0 – 3, respectively, and then sums these scores over all questions, with responses to positively worded questions reverse-coded. The WLS modifies this scoring methodology and simply sums the number of days reported over all 20 questions, with positively-worded questions reverse-coded. Responses to positive questions were reverse coded. The resulting measure of depression ranges from 0 to 140.

The Spielberger Anxiety Index and Anger Index scores are derived from a subset of questions from the State-Trait Anxiety Inventory (STAI) [7] and the State-Trait Anger Expression Inventory-II (STAXI-2) [8].STAI consists of 40 questions – 20 meant to measure state anxiety and 20 to measure trait anxiety – and the STAXI-2 consists of 57 questions – 15 to measure state anger, 10 to measure trait anger, and 27 to measure anger expression. The responses are generally measured on a four point Likert scale. The version administered by the WLS is on a seven point scale measuring the number of days in the past week that the subject had the particular feeling addressed by the question. The Spielberger Anxiety and Anger Index questionnaire administered by the WLS included seven questions from the state anxiety and state anger inventories, respectively. Both scores were on a 49 point scale. See pages 136-137 in http://www.ssc.wisc.edu/wlsresearch/documentation/scales/WlsScalesDoc_Nov2010.pdf for detailed description of these measures.

[Spielberger, C. D. (1999). State-Trait Anger Expression Inventory-2. Lutz, Florida: Psychological Assessment Resources
Spielberger, C. D. (1999). State-Trait Anger Expression Inventory-2. Lutz, Florida: Psychological Assessment Resources

Outcome Availability: Our two primary outcomes were based on the age 65 Letter Fluency (LF) score, Delayed Word Recall (DWR) score, and modified CES-D score. These outcomes were not always available for each subject. eTable 2 shows the availability of primary outcome components.

Both primary outcomes were unavailable for 1,212 (31.05%) of our subjects and they were subsequently excluded from the primary analyses leaving us with a sample size of 2,692. The mixed effects models fit as secondary analyses included many of the 1,212 subjects excluded from the primary analysis.

Other outcomes: Subjects were classified as heavy drinkers if they reported having five or more drinks on five or more occasions in the month of their WLS interview.

Availability of data and analysis code to researchers

The data used in the analyses contains some variables available in the public release of the Wisconsin Longitudinal Study and some variables only available in the protected data release of the Wisconsin Longitudinal Study. To access the protected data contact the WLS at wls@ssc.wisc.edu. Researchers who obtain access to the WLS protected data will also be given access to the files and code we used in these analyses on request.

Supplemental statistical details

Details on Matching

© 2017 American Medical Association. All rights reserved.

A common strategy in observational studies is to create matched sets consisting of exposed and control subjects who are comparable. Ideally, within a matched set, exposed and control subjects would be identical along these baseline variables. With several baseline variables, however, it is typically impossible to achieve this goal with any reasonable sample size. Instead, we aim to create matched sets which *balance* the distributions of each of the baseline variables between exposed and control groups.

Since age 65 LW, DWR, and CES-D scores were not always available for every subject, we first stratify our sample based on the availability of these scores (i.e. according to the rows of Table 2). Within each stratum, we use optimal full matching to construct matched sets of exposed and control subjects. Full matching returns matched sets with one exposed and one or more control subjects or one control and one or more exposed subjects. The sub-classification induced by an unconstrained, full match is optimal for an observational study[9] and does not discard any observations, unlike pair matching. However, treatment effect estimates can have poor precision relative to optimal pair matching if some matched sets are very large. To address this, we constrain the ratio of exposed to control subjects within any matched set to be between 1:6 and 6:1 to avoid matchings that contain extremely large matched sets which tend to decrease the precision of treatment effect estimates. In practice, the increase in bias along covariates due to this constraint tends to be small relative to the improvement in variance reduction[10]. To achieve covariate balance, our full matching procedure first uses a propensity score caliper[11] to ensure that matched subjects are close on estimated propensity score and then minimizes the rank-based Mahalanobis distance between covariates within propensity score calipers. The rank-based Mahalanobis distance is insensitive to outliers and takes into account the correlation between covariates[10]. Propensity scores have the property of balancing covariates between treated and control groups when conditioned upon. The use of a propensity score caliper takes advantage of this property to achieve a certain level of balance before explicitly matching on covariates[12].

The resulting matched sets consist of either one to six football player and a single control or a single football player and one to six control subjects. To assess balance, we look at the standardized difference, which is the average over the matched sets of the mean difference between the exposed and control subjects in the matched set as a fraction of the within group pre-matching standard deviation (specifically the square root of the average of the variance among exposed subjects and the variance among control subjects). After matching, we use regression-based covariate adjustment to remove any residual covariate imbalance. In matching, we aim to achieve standardized differences below 0.2 SDs for all baseline variables, as biases due to imbalances of this magnitude may be removed with model-based covariate adjustment[13,14]. Tables S3 and S4 are analogous to Table 1 and show pre- and post-matching standardized differences when matching football players to our two alternative control groups: non sport controls (eTable 3) and non-collision sport controls (eTable 4). eTable 5 shows the standardized differences from matching non-collision sport and non-sport controls.

We note that some subjects are missing measurements of certain baseline variables. To deal with missing data in matched observational studies, we follow the recommendation of Rosenbaum and Rubin (1984)[15]: for each baseline variable we create an indicator variable of missingness. We include these indicators in the matching algorithm along with the baseline covariates. When assessing the suitability of a given match, we also consider the balance in the proportion of missingness of each variable. See Chapter 9.4 of Rosenbaum (2002)[10] for further details. Tables S6-S9 compare the balance of missingness indicators before and after matching for each of the four matches.

Our matching procedure results in matched sets consisting of either one or more football player and a single control or a single football player and one or more control subjects. eTable 1 shows the composition of the matched sets for each of the four matches.

Graphical Displays

Tables S3 – S9 assessed balance by looking only at the standardized differences. This notion of balance requires only that the mean of the treated subjects and weighted mean of matched controls be similar. This requirement is weaker than requiring that the distributions themselves are similar. We use a weighted variation of letter-valued box plots[16] (Hofmann et al., 2006) to assess the similarities in distributions between exposed and control groups after

© 2017 American Medical Association. All rights reserved.

matching. Standard boxplots show the median, interquartile range (IQR), whiskers to denote quartiles to "adjacent values,"[17] and outliers. In large datasets, however, the number of labeled outliers in the standard boxplot can grow quickly and obscure useful information about the tails of the distribution that otherwise could be described accurately when samples are large. Letter-valued box plots were introduced to address these shortcomings. The k-th letter values correspond to the tail areas of size 2^{-k} [18]. In addition to the IQR box, boxes with upper and lower limits determined by letter values are displayed to convey information about the shape and spread of the distribution toward the tails. Outliers are defined as observations outside of the most extreme letter-valued box. The depth of the letter-valued boxes to be displayed can be determined by the size of the sample so that the quantiles they represent can be precisely estimated. The weighted variation we introduce constructs the letter values using the weighted sample arising from the full matching procedure.

The matched sample returned by the full matching procedure is weighted to the treatment group in order to assess the average treatment effect on the treated (ATT) – the long-term cognitive and psychological effect of playing high school football on students who played high school football. Arguably, this is the quantity of interest when studying the public health consequences of high school football participation. In such a weighting scheme, treated units are given weight one and control units in a given matched set have weight equal to the number of treated units in the set divided by the number of control units in the set.

Ordered Hypothesis Testing

We now detail the ordered hypothesis testing procedure and the systematic variation of the control groups. Students who participate in high school sports may differ substantially from non-participants in terms of personality, temperament, and overall fitness and lifestyle (all of which are unmeasured). The non-collision sport control group is arguably a closer and more appropriate control group for our study and including non-sport controls may introduce problematic unmeasured confounding. However, simply dropping the non-sport controls from our analysis would cut our effective sample size by about 1/3 and may result in a substantial decrease in power. By constructing comparisons between the treated group and all controls as well as both control groups separately we systematically vary the unmeasured confounders of concern. Agreement across these three comparisons provide evidence that an ostensible treatment effect is a result of playing football and not due to unmeasured differences in students who played sports in high school and those who did not. Equivalence between the non-collision sport and non-sport control groups would provide further evidence that we are indeed testing for the effect of playing high school football.

In order to preserve the increased power of using controls from both groups while still testing the treated group against each group separately we follow an ordered testing procedure which controls the family wise error rate (FWER)[19] In particular, we first test the null of no treatment effect using matched sets constructed with all controls. If we reject that at level α, we conduct the same test separately using matched sets constructed using non collision-sport controls and non-sport controls. If we reject both separate tests at level α, we perform an equivalence test between the two control groups. If at any stage of the ordered testing procedure we do not reject, we stop the procedure. For example, if we do not reject the test using all controls we do not continue on to test against the two control groups separately. This stopping rule guarantees FWER control at level α.

We run this procedure for each of our primary outcomes at level α=0.025 using Bonferroni-Holm correction to control overall FWER for our primary analyses at level α=0.05. For completeness, we also report marginal 97.5% confidence intervals that go along with each test, regardless of whether we actually reached the test in the ordered testing procedure. For each outcome variable, we judge effect sizes by Cohen's popular criterion – 0.2, 0.5, and 0.8 SDs for small, medium, and large effects, respectively[20]. For the combined cognition score, these cut-offs are 0.16, 0.41, and 0.65. For the CES-D score, the thresholds are 2.56, 6.41, and 10.25.

Primary and secondary analysis restricted to subjects with *ApoE* e4 data

© 2017 American Medical Association. All rights reserved.

The WLS undertook a genotyping study starting in 2007[2]. We focus on SNP data for rs429358 and rs7412 and counted the number of e4 variants. In all, genotype data was missing for 2,224 (56.97%) of the initial 3,904 eligible subjects. Of the 2,692 subjects included in our analysis, however, *ApoE* data was missing from only 1,084 subjects (40.27%). eTable 11 shows the availability of primary outcome components among the remaining 1,680 subjects with available *ApoE* data. Of these subjects, 72 (4.48%) were missing LF, DWR, and CES-D scores and were dropped from our analysis.

We repeat our primary analysis with the remaining 1,608 subjects. Once again, we stratify the subjects based on the available primary outcome components and perform full matching within each stratum. We do this for each of the four comparisons – football vs all controls, football vs non-sport controls, football vs non-collision sport controls, and non-sport vs non-collision sport controls. This time, however, we include the number of e4 variants in the matching procedure. Tables S12 – S15 shows the balance of our matched sets on the covariates, eTable 16-S19 shows the balance of the matched sets on the missingness of the covariates and Tables S20 show the composition of the matched sets for this analysis.

eTable 21 is the analog of Table 2 in the main text: it shows the estimated effect and 97.5% confidence intervals from each comparison. Among subject with available *ApoE* genotype data, we find that football players had slightly lower combined cognition scores than all matched controls but the different was not significant ($p = 0.07$), though the confidence interval does contain some small-sized harmful effects. Compared to both alternative control groups, we have similar findings.

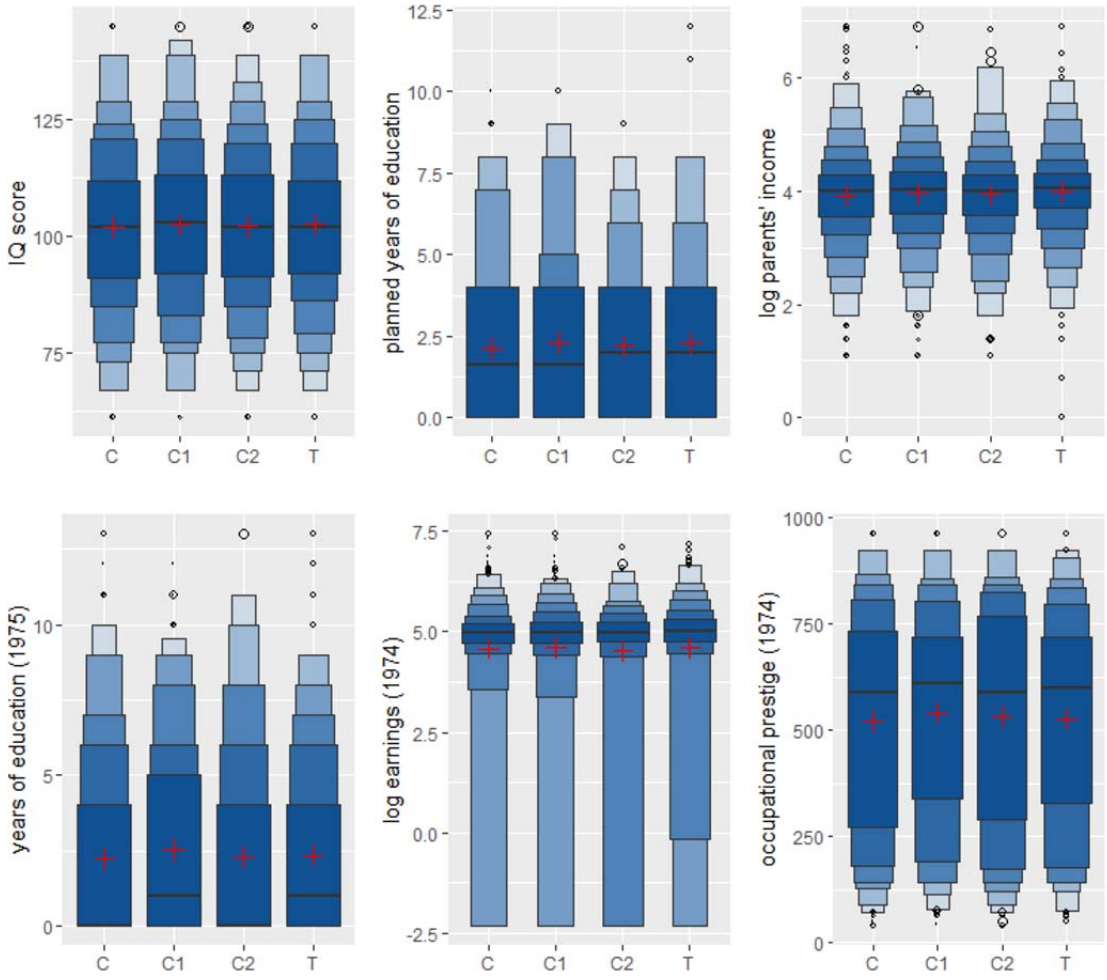
Football players had slightly higher modified CES-D scores but this was not significant ($p = 0.94$) and the confidence interval contained only values smaller than the small effect size cut-off. Compared to non-collision sport controls, football players had lower CES-D scores. This difference is marginally significant ($p = 0.02$) but this is not significant within the ordered hypothesis testing framework. The corresponding confidence interval does contain values corresponding to small effects but we observe that they are all negative, indicating football players reported less depressive symptoms than non-collision sport controls. In short, across all comparisons, we do not find evidence of a significant harmful effect of playing football on depression.

eTable 22 shows the effect of playing football on the secondary outcomes. We see that these effects are all insignificant. We do find that some confidence intervals contain some small-sized and medium-sized effects, but this is not particularly surprising, given the reduced sample size in this restricted analysis. When estimating the effect of playing high school football on heavy drinking status at ages 54 and 72, the conditional logit model failed to converge, yielding a numerically unstable estimated treatment effect.

[2] Further details may be found at the WLS website:
http://www.ssc.wisc.edu/wlsresearch/WLS_CIDR_Summary.pdf

© 2017 American Medical Association. All rights reserved.

eFigure. Comparison of Post-Matching Distributions of Important Baseline Covariates and Post-Exposure Variables by Treatment and Control Group



eFigure: The top panel compares the post-matching distributions of important covariates for football players (T), all controls (C), non-sport controls (C1), and other-sport controls (C2). The bottom panel compares the post-matching distributions of important post-exposure variables. The lower limits of the boxes indicate the weighted, empirical 2nd, 3rd, 6th, 13th and 25th percentiles (going from bottom to top) and the upper limits indicate the 98th, 97th 94th, 87th, and 75th percentiles (going from top to bottom) where observations are weighted according to the composition of matched sets (see eTable 1 in Supplement). The sizes of the outliers are proportional to their weights. The red cross indicates the weighted sample mean and the black horizontal line indicates the weighted sample median.

© 2017 American Medical Association. All rights reserved.

eTable 1. Composition of Matched Sets

Composition	Football – All Controls	Football & Non-Sport Controls	Football & Non-Collision sport controls	Non-Sport & Non-Collision Sport Controls
6:1	0	0	0	1
5:1	0	0	0	2
4:1	0	0	0	6
3:1	0	97	125	6
2:1	0	17	60	20
1:1	550	315	318	259
1:2	52	49	10	33
1:3	41	16	5	19
1:4	25	20	3	27
1:5	15	17	2	13
1:6	151	92	1	117

eTable 1: Composition of Matched Sets for all 4 matches. For matches involving football players, the first number in the Composition column records the number of football players and the second records the number of controls in the matched set. For the match involving both control groups, the first number records the number of non-sport controls and the second records the number of non-collision sport controls in the matched set.

© 2017 American Medical Association. All rights reserved.

eTable 2. Sports Played by Eligible Participants

Sport	Number	Sport	Number
Baseball	577	Lacrosse	0
Basketball	845	Soccer	0
Cross Country	68	Swimming	49
Curling	32	Track	619
Football	1153	Tennis	55
Gymnastics	0	Volleyball	97
Hockey	11	Wrestling	114

eTable 2: Number of eligible participants who participated in each sport

© 2017 American Medical Association. All rights reserved.

eTable 3: Availability of Letter Fluency, Delayed Word Recall, and Modified CES-D Scores at Age 65 Years

Available Scores	Football	Non-Sport Control	Non-Collision Sport Controls	All Controls
LF, DWR, CES-D	467	682	301	983
LF, DWR	56	118	40	158
LF, CES-D	24	37	19	56
DWR, CES-D	55	92	33	125
LF	9	17	4	21
DWR	13	14	10	24
CES-D	210	332	159	491
None	319	659	234	893

eTable 3: Availability of Primary Outcomes for exposed and control subjects

© 2017 American Medical Association. All rights reserved.

eTable 4. Mean Baseline Covariates Before and After Matching Football Players to Non-Sport Controls (Analogous to Table 1 in Main Text)

	Football Players	Non-sport Controls (Before Matching)	Non-sport Controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957	577.67	532.92	574.25	0.19	0.01
High school size (N)	134.86	205.51	139.99	-0.57	-0.04
High school rank (quantile)	48.81	44.12	49.17	0.17	-0.01
Parental Income in 1957 ($100)	69.38	62.93	66.97	0.10	0.04
Father's Education (years)	9.91	9.60	9.91	0.09	0.00
Mother's Education (years)	10.80	10.39	10.72	0.15	0.03
Duncan Socioeconomic Index of father's job in 1957	355.30	342.90	352.13	0.05	0.01
Planned future education (years)	2.34	1.66	2.36	0.32	-0.01
IQ	102.39	101.60	102.44	0.05	0.00
Considered an "outstanding student" by teacher (%)	12	10	13	0.06	-0.04
Participated in band, orchestra, chorus, or musical ensembles (%)	34	26	35	0.16	-0.02
Participated in drama, speech, debate (%)	29	18	29	0.27	0.00
Participated in school government (%)	36	10	31	0.65	0.11
Participated in school publications (%)	23	13	23	0.27	0.00
Father was a farmer (%)	18	18	18	0.02	-0.01
Planned to serve in military (%)	27	26	24	0.01	0.07
Attended Catholic high school (%)	7	14	0.09	-0.25	-0.09
Lived with both parents (%)	91	91	91	0.01	0.02
Mother working in 1957 (%)	37	36	36	0.02	0.02
Teachers encouraged college (%)	58	42	56	0.33	0.03
Parents encouraged college (%)	0.65	0.58	0.67	0.16	-0.05
Had friends who	43	34	46	0.19	-0.07

© 2017 American Medical Association. All rights reserved.

planned to go to college (%)					
Extent to which subject discussed future plans with teachers					
Not at all (%)	24	33	25	-0.16	-0.01
Sometimes (%)	66	60	64	0.10	0.03
Very much (%)	10	7	11	0.06	-0.02
Extent to which subject discussed future plans with parents					
Not at all (%)	2	2	1	0.00	0.01
Sometimes (%)	42	47	43	-0.09	-0.02
Very much (%)	56	51	55	0.09	0.01
Family wealth relative to community					
Considerably below average (%)	0	0	0	0.00	0.00
Somewhat below average (%)	5	6	6	-0.02	-0.01
Average (%)	70	70	68	0.00	0.03
Somewhat above average (%)	22	21	24	0.00	-0.04
Considerably above average	3	1	2	0.02	0.02
Financial support from parents for college					
Cannot support (%)	28	32	28	-0.07	0.00
Can support, with some sacrifice (%)	58	55	60	0.05	-0.03
Easily support (%)	14	13	12	0.03	0.04
Number of *ApoE* e4 variants	0.30	0.28	0.27	0.05	0.07

eTable 4: Average baseline covariates for football players and non-sport controls, before and after matching. Before matching control values are unweighted and after matching control values are weighted according to the composition of the matched sets (see eTable 1). We did not explicitly match on the number ApoE *e4 variants but we nevertheless assess the balance of this variable before and after matching.*

© 2017 American Medical Association. All rights reserved.

eTable 5. Mean Baseline Covariates Before and After Matching Football Players to Non-Collision Sport Controls (Analogous to Table 1 in Main Text)

	Football Players	Non-collision sport controls (Before Matching)	Non-collision sport Controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957	577.67	562.24	565.33	0.07	0.05
High school size (N)	134.86	148.43	131.50	-0.12	0.03
High school rank (quantile)	48.81	47.34	48.35	0.05	0.02
Parental Income in 1957 ($100)	69.38	61.62	64.68	0.11	0.07
Father's Education (years)	9.91	10.12	9.96	-0.06	-0.01
Mother's Education (years)	10.80	10.75	10.73	0.02	0.02
Duncan Socioeconomic Index of father's job in 1957	355.30	332.89	348.54	0.10	0.03
Planned future education (years)	2.34	2.08	2.24	0.12	0.05
IQ	102.39	101.58	102.14	0.05	0.02
Considered an "outstanding student" by teacher (%)	12	12	12	-0.02	-0.01
Participated in band, orchestra, chorus, or musical ensembles (%)	34	33	31	0.01	0.05
Participated in drama, speech, debate (%)	29	25	28	0.09	0.01
Participated in school government (%)	36	27	34	0.19	0.04
Participated in school publications (%)	23	23	22	0.00	0.03
Father was a farmer (%)	18	21	16	-0.06	0.05
Planned to serve in military (%)	27	29	28	-0.05	-0.03
Attended Catholic high school (%)	7	4	6	0.10	0.03
Lived with both parents (%)	91	90	90	0.04	0.03
Mother working in 1957 (%)	37	36	34	0.02	0.07
Teachers encouraged college (%)	58	53	57	0.09	0.01
Parents encouraged college (%)	65	66	66	-0.02	-0.01
Had friends who planned to go to	43	42	42	0.02	0.02

© 2017 American Medical Association. All rights reserved.

college (%)					
Extent to which subject discussed future plans with teachers					
Not at all	24	26	25	-0.03	-0.01
Sometimes	66	64	66	0.03	-0.01
Very much	10	10	9	0.00	0.02
Extent to which subject discussed future plans with parents					
Not at all	2	3	3	-0.02	-0.01
Sometimes	42	42	43	-0.01	-0.01
Very much	56	54	55	0.03	0.03
Family wealth relative to community					
Considerably below average	0.00	1	0	0.00	0.02
Somewhat below average	5	9	7	-0.06	-0.04
Average	70	70	69	0.00	0.02
Somewhat above average	22	18	21	0.05	0.01
Considerably above average	3	2	3	0.01	0.00
Financial support from parents for college					
Cannot support (%)	28	34	33	-0.09	-0.08
Can support, with some sacrifice (%)	58	53	54	0.08	0.07
Easily support (%)	14	13	13	0.02	0.01
Number of *ApoE* e4 variants	0.30	0.25	0.25	0.13	0.14

eTable 5: Average baseline covariates for football players and non-collision sport controls, before and after matching. Before matching control values are unweighted and after matching control values are weighted according to the composition of the matched sets (see eTable 1). We did not explicitly match on the number ApoE *e4 variants but we nevertheless assess the balance of this variable before and after matching.*

© 2017 American Medical Association. All rights reserved.

eTable 6. Mean Baseline Covariates Before and After Matching Non-Collision Sport Controls to Non-Sport Controls (Analogous to Table 1 in Main Text)

	Non-collision sport control	Non-sport controls (Before Matching)	Non-sport controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957	562.24	532.92	612.29	0.13	-0.22
High school size (N)	148.43	205.51	157.57	-0.44	-0.07
High school rank (quantile)	47.34	44.12	49.82	0.12	-0.09
Parental Income in 1957 ($100)	61.62	62.93	63.54	-0.02	-0.03
Father's Education (years)	10.12	9.60	10.20	0.15	-0.02
Mother's Education (years)	10.75	10.39	10.92	0.13	-0.06
Duncan Socioeconomic Index of father's job in 1957	332.89	342.90	346.06	-0.04	-0.06
Planned future education (years)	2.08	1.66	2.23	0.20	-0.07
IQ	101.58	101.60	103.05	0.00	-0.10
Considered an "outstanding student" by teacher (%)	12	10	12	0.08	0.00
Participated in band, orchestra, chorus, or musical ensembles (%)	33	26	35	0.15	-0.03
Participated in drama, speech, debate (%)	25	18	30	0.18	-0.11
Participated in school government (%)	27	10	26	0.45	0.01
Participated in school publications (%)	23	13	24	0.27	-0.03
Father was a farmer (%)	21	18	21	0.08	0.00
Planned to serve in military (%)	29	26	24	0.06	0.12
Attended Catholic high school (%)	4	14	7	-0.34	-0.10
Lived with both parents (%)	90	91	93	-0.03	-0.08
Mother working in 1957 (%)	36	36	36	-0.01	0.00
Teachers encouraged college (%)	53	42	55	0.24	-0.03
Parents encouraged college (%)	66	58	70	0.17	-0.08
Had friends who planned to go to college (%)	42	34	46	0.17	-0.11

© 2017 American Medical Association. All rights reserved.

Extent to which subject discussed future plans with teachers					
Not at all	26	33	25	-0.13	0.01
Sometimes	64	60	64	0.07	0.00
Very much	10	7	11	0.05	-0.01
Extent to which subject discussed future plans with parents					
Not at all	3	2	1	0.02	0.03
Sometimes	42	47	43	-0.08	-0.01
Very much	54	51	56	0.07	-0.02
Family wealth relative to community					
Considerably below average	1	0	0	0.00	0.00
Somewhat below average	9	6	6	0.04	0.05
Average	70	70	69	0.00	0.02
Somewhat above average	18	21	23	-0.05	-0.09
Considerably above average	2	1	1	0.01	0.01
Financial support from parents for college					
Cannot support (%)	34	32	26	0.03	0.12
Can support, with some sacrifice (%)	53	55	61	-0.03	-0.13
Easily support (%)	13	13	12	0.00	0.01
Number of *ApoE* e4 variants	0.25	0.28	0.25	-0.08	0.00

eTable 6: Average baseline covariates for non-collision sport controls and non-sport controls. Before matching non-sport control values are unweighted and after matching non-sport control values are weighted according to the composition of the matched sets (see eTable 1). We did not explicitly match on the number ApoE *e4 variants but we nevertheless assess the balance of this variable before and after matching.*

© 2017 American Medical Association. All rights reserved.

eTable 7. Percentages of Missing Baseline Data Before and After Matching Football Players to All Controls

	Football Players	All Controls (Before Matching)	All Controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957 (%)	35	33	34	0.03	0.01
High school rank (%)	4	7	5	-0.11	-0.03
Parental Income in 1957 (%)	3	4	3	-0.05	-0.01
Father's Education (years)	1	3	2	-0.09	-0.03
Mother's Education (years)	1	1	1	-0.07	-0.05
Duncan Socioeconomic Index of father's job in 1957	1	1	1	-0.03	-0.02
Planned future education (years)	8	8	7	-0.01	0.03
Lived with both parents (%)	0	0	0	0.05	0.06
Mother working in 1957 (%)	0	1	0	-0.03	0.01
Teachers encouraged college (%)	4	4	4	-0.04	-0.04
Parents encouraged college (%)	2	4	3	-0.08	-0.04
Had friends who planned to go to college (%)	12	13	11	-0.03	0.03
Extent to which subject discussed future plans with teachers	2	3	2	-0.03	-0.01
Extent to which subject discussed future plans with parents	1	1	1	0.01	0.01
Family wealth relative to community	2	4	2	-0.04	0.01
Financial support from parents for college	12	14	12	-0.05	0.01
ApoE e4	38	41	39	-0.06	-0.03

eTable 7: Percentage of football players and all controls missing baseline covariate values, before and after matching. Before matching control values are unweighted and after matching control values are weighted according to the composition of the matched set (eTable 1).

© 2017 American Medical Association. All rights reserved.

eTable 8. Percentages of Missing Baseline Data Before and After Matching Football Players to Non-Sport Controls

	Football Players	Non-sport Controls (Before Matching)	Non-sport Controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957 (%)	35	33	32	0.03	0.05
High school rank (%)	4	7	4	-0.14	-0.01
Parental Income in 1957 (%)	3	4	3	-0.06	0.03
Father's Education (years)	1	3	1	-0.12	0.01
Mother's Education (years)	1	1	1	-0.06	0.00
Duncan Socioeconomic Index of father's job in 1957	1	1	1	0.00	-0.02
Planned future education (years)	8	9	8	-0.04	0.01
Lived with both parents (%)	0	0	0	0.04	0.06
Mother working in 1957 (%)	0	0	0	0.00	0.04
Teachers encouraged college (%)	4	4	3	-0.02	-0.03
Parents encouraged college (%)	2	4	3	-0.08	-0.02
Had friends who planned to go to college (%)	12	15	12	-0.07	0.00
Extent to which subject discussed future plans with teachers	2	2	2	-0.01	-0.01
Extent to which subject discussed future plans with parents	1	1	1	0.01	0.00
Family wealth relative to community	2	4	2	-0.04	0.00
Financial support from parents for college	12	14	11	-0.06	0.02
Number of *ApoE* e4 variants	38	42	38	-0.07	0.01

eTable 8: Percentage of football players and non-sport controls missing baseline covariate values, before and after matching. Before matching non-sport control values are unweighted and after matching non-sport control values are weighted according to the composition of the matched set (eTable 1).

© 2017 American Medical Association. All rights reserved.

eTable 9. Percentages of Missing Baseline Data Before and After Matching Football Players to Non-Collision Sport Controls

	Football Players	Non-collision sport-Controls (Before Matching)	Non-collision sport-Controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957 (%)	35	32	33	0.05	0.03
High school rank (%)	4	5	4	-0.06	0.00
Parental Income in 1957 (%)	3	4	3	-0.04	-0.02
Father's Education (years)	1	2	1	-0.02	0.00
Mother's Education (years)	1	2	1	-0.1	0.01
Duncan Socioeconomic Index of father's job in 1957	1	1	1	-0.08	-0.04
Planned future education (years)	8	6	6	0.07	0.07
Lived with both parents (%)	0	0	0	0.07	0.07
Mother working in 1957 (%)	0	1	0	-0.08	-0.01
Teachers encouraged college (%)	4	5	3	-0.07	0.02
Parents encouraged college (%)	2	4	2	-0.08	0.04
Had friends who planned to go to college (%)	12	10	11	0.07	0.04
Extent to which subject discussed future plans with teachers	2	4	2	-0.08	0.00
Extent to which subject discussed future plans with parents	1	1	1	0.00	0.02
Family wealth relative to community	2	4	2	-0.04	0.02
Financial support from parents for college	12	13	12	-0.03	0.01
Number of *ApoE* e4 variants	38	40	41	-0.04	-0.05

eTable 9: Percentage of football players and non-collision sport controls missing baseline covariate values, before and after matching. Before matching non-collision sport control values are unweighted and after matching control values are weighted according to the composition of the matched set (eTable 1).

© 2017 American Medical Association. All rights reserved.

eTable 10. Percentages of Missing Baseline Data Before and After Matching Non-Collision Sport Controls to Non-Sport Controls

	Non-Collision Sport Controls	Non-Sport Controls (Before Matching)	Non-Sport Controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957 (%)	32	33	30	-0.02	0.04
High school rank (%)	5	7	5	-0.07	0.03
Parental Income in 1957 (%)	4	4	3	-0.01	0.06
Father's Education (years)	2	3	2	-0.10	-0.01
Mother's Education (years)	2	1	1	0.04	0.05
Duncan Socioeconomic Index of father's job in 1957	1	1	1	0.08	0.05
Planned future education (years)	6	9	8	-0.10	-0.06
Lived with both parents (%)	0	0	0	-0.04	-0.01
Mother working in 1957 (%)	1	0	0	0.08	0.09
Teachers encouraged college (%)	5	4	4	0.05	0.08
Parents encouraged college (%)	4	4	3	0.00	0.04
Had friends who planned to go to college (%)	10	15	12	-0.14	-0.06
Extent to which subject discussed future plans with teachers	4	2	3	0.06	0.02
Extent to which subject discussed future plans with parents	1	1	1	0.00	-0.01
Family wealth relative to community	4	4	3	-0.01	0.01
Financial support from parents for college	13	14	11	-0.03	0.07
Number of *ApoE* e4 variants	40	42	40	-0.03	0.01

eTable 10: Percentage of non-collision sport controls and non-sport controls missing baseline covariate values, before and after matching. Before matching non-sport control values are unweighted and after matching non-sport control values are weighted according to the composition of the matched set (eTable 1).

© 2017 American Medical Association. All rights reserved.

eTable 11. Analog to eTable 1, Restricted to Participants With Available *ApoE* Genotype Data

Composition	Football – All Controls	Football & Non-Sport Controls	Football & Non-Collision sport controls	Non-Sport & Non-Collision Sport Controls
6:1	0	0	0	2
5:1	0	0	0	4
4:1	0	0	0	0
3:1	0	63	87	4
2:1	0	9	25	11
1:1	349	206	184	149
1:2	39	16	2	29
1:3	23	20	5	11
1:4	12	7	0	7
1:5	15	9	1	6
1:6	78	51	3	70

eTable 11: Analog of eTable 1, but restricted to subjects with available ApoE *genotype data.*

© 2017 American Medical Association. All rights reserved.

eTable 12. Analog of eTable 3, Restricted to Participants With Available *ApoE* Genotype Data

Available Scores	Football	Non-Sport Control	Non-Collision Sport Controls	All Controls
LF, DWR, CES-D	323	453	215	668
LF, DWR	23	42	11	53
LF, CES-D	15	19	3	22
DWR, CES-D	34	46	16	62
LF	0	4	1	5
DWR	5	5	4	9
CES-D	116	184	89	273
None	21	27	24	51

eTable 12: Analog to eTable 3, but restricted only to those subjects for whom ApoE *genotype data was available.*

© 2017 American Medical Association. All rights reserved.

eTable 13. Analog of Table 1 in Main Text, Restricted to Participants With Available *ApoE* Genotype Data

	Football Players	All Controls (Before Matching)	All Controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957	580.09	556.40	563.96	0.10	0.07
High school size (N)	131.43	186.38	138.85	-0.45	-0.06
High school rank (quantile)	49.91	46.84	48.62	0.11	0.05
Parental Income in 1957 ($100)	69.33	62.59	66.55	0.10	0.04
Father's Education (years)	9.94	9.89	9.88	0.01	0.02
Mother's Education (years)	10.88	10.59	10.79	0.10	0.04
Duncan Socioeconomic Index of father's job in 1957	349.69	338.65	341.98	0.05	0.03
Planned future education (years)	2.41	1.85	2.28	0.26	0.06
IQ	103.04	102.65	102.81	0.03	0.02
Considered an "outstanding student" by teacher (%)	12	12	12	0.02	0.02
Participated in band, orchestra, chorus, or musical ensembles (%)	34	28	32	0.13	0.06
Participated in drama, speech, debate (%)	30	22	27	0.17	0.06
Participated in school government (%)	37	15	28	0.51	0.20
Participated in school publications (%)	25	17	22	0.20	0.07
Father was a farmer (%)	21	20	20	0.02	0.02
Planned to serve in military (%)	28	28	29	0.02	-0.01
Attended Catholic high school (%)	6	10	9	-0.16	-0.09
Lived with both parents (%)	91	90	91	0.04	0.02
Mother working in 1957 (%)	37	36	37	0.02	0.00
Teachers encouraged college (%)	61	47	56	0.29	0.10
Parents encouraged college (%)	67	62	66	0.11	0.02
Had friends who planned to go to college (%)	44	38	44	0.12	0.00

© 2017 American Medical Association. All rights reserved.

Extent to which subject discussed future plans with teachers					
Not at all	25	29	26	-0.07	-0.01
Sometimes	63	63	64	0.01	-0.01
Very much	12	8	10	0.06	0.03
Extent to which subject discussed future plans with parents					
Not at all	3	2	2	0.01	0.01
Sometimes	42	46	44	-0.07	-0.03
Very much	55	52	54	0.05	0.02
Family wealth relative to community					
Considerably below average	1	0	0	0.00	0.00
Somewhat below average	6	7	7	-0.02	-0.02
Average	69	69	67	0.02	0.04
Somewhat above average	22	22	23	-0.01	-0.03
Considerably above average	3	2	2	0.01	0.01
Financial support from parents for college					
Cannot support (%)	27	32	29	-0.08	-0.02
Can support, with some sacrifice (%)	59	55	58	0.07	0.03
Easily support (%)	14	13	14	0.01	0.00
Number of *ApoE* e4 variants	0.30	0.27	0.29	0.06	0.02

eTable 13: Analog to Table 1 in main text, but restricted to subjects with available ApoE *genotype data. The weighting of after matching values for all controls is determined by the matched set composition (eTable 11).*

© 2017 American Medical Association. All rights reserved.

eTable 14. Analog of eTable 4, Restricted to Participants With Available ApoE Genotype Data

	Football Players	Non-Sport controls (Before Matching)	Non-sport controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957	580.09	546.16	564.82	0.14	0.07
High school size (N)	131.43	203.43	138.30	-0.59	-0.06
High school rank (quantile)	49.91	45.77	49.90	0.15	0.00
Parental Income in 1957 ($100)	69.33	62.70	69.08	0.10	0.00
Father's Education (years)	9.94	9.70	9.97	0.07	-0.01
Mother's Education (years)	10.88	10.50	10.88	0.14	0.00
Duncan Socioeconomic Index of father's job in 1957	349.69	342.83	342.58	0.03	0.03
Planned future education (years)	2.41	1.72	2.40	0.33	0.01
IQ	103.04	102.63	102.91	0.03	0.01
Considered an "outstanding student" by teacher (%)	12	11	12	0.05	0.00
Participated in band, orchestra, chorus, or musical ensembles (%)	34	26	31	0.18	0.08
Participated in drama, speech, debate (%)	30	21	30	0.21	0.00
Participated in school government (%)	37	10	30	0.67	0.18
Participated in school publications (%)	25	13	22	0.31	0.09
Father was a farmer (%)	21	19	22	0.05	-0.04
Planned to serve in military (%)	28	27	26	0.03	0.05
Attended Catholic high school (%)	6	13	8	-0.24	-0.06
Lived with both parents (%)	91	90	92	0.04	-0.01
Mother working in 1957 (%)	37	37	37	0.01	-0.01
Teachers encouraged college (%)	61	44	59	0.36	0.04
Parents encouraged college (%)	67	60	65	0.15	0.03
Had friends who planned to go to college (%)	44	35	43	0.19	0.03

© 2017 American Medical Association. All rights reserved.

Extent to which subject discussed future plans with teachers					
Not at all	25	31	25	-0.10	0.00
Sometimes	63	62	64	0.03	-0.01
Very much	12	7	11	0.07	0.01
Extent to which subject discussed future plans with parents					
Not at all	3	2	2	0.02	0.02
Sometimes	42	47	45	-0.09	-0.05
Very much	55	51	53	0.07	0.03
Family wealth relative to community					
Considerably below average	1	0	0	0.01	0.01
Somewhat below average	6	7	7	-0.02	-0.02
Average	69	68	68	0.02	0.03
Somewhat above average	22	23	23	-0.02	-0.02
Considerably above average	3	2	2	0.02	0.01
Financial support from parents for college					
Cannot support (%)	27	32	26	-0.08	0.02
Can support, with some sacrifice (%)	59	55	61	0.07	-0.03
Easily support (%)	14	13	13	0.01	0.01
Number of ApoE e4 variants	0.30	0.28	0.28	0.04	0.04

eTable 14: Analog to eTable 4, but restricted to subjects with available ApoE *genotype data. The weighting of after matching values for non-sport controls is determined by the matched set composition (eTable 11).*

© 2017 American Medical Association. All rights reserved.

eTable 15. Analog of eTable 5, Restricted to Participants With Available ApoE Genotype Data

	Football Players	Non-Collision Sport Controls (Before Matching)	Non-Collision Sport (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957	582.26	578.42	568.13	0.02	0.06
High school size (N)	131.50	148.62	130.51	-0.15	0.01
High school rank (quantile)	50.14	49.22	50.19	0.03	0.00
Parental Income in 1957 ($100)	69.28	62.33	69.50	0.09	0.00
Father's Education (years)	9.94	10.31	10.02	-0.10	-0.02
Mother's Education (years)	10.84	10.80	10.85	0.02	0.00
Duncan Socioeconomic Index of father's job in 1957	350.23	329.25	341.51	0.09	0.04
Planned future education (years)	2.41	2.11	2.27	0.14	0.06
IQ[3]	103.01	102.68	103.33	0.02	-0.02
Considered an "outstanding student" by teacher (%)	13	13	12	-0.02	0.03
Participated in band, orchestra, chorus, or musical ensembles (%)	35	34	33	0.03	0.04
Participated in drama, speech, debate (%)	30	26	26	0.09	0.08
Participated in school government (%)	37	27	34	0.21	0.05
Participated in school publications (%)	25	25	24	0.01	0.04
Father was a farmer (%)	21	22	21	-0.02	0.01
Planned to serve in military (%)	28	28	28	-0.01	-0.01
Attended Catholic high school (%)	6	4	6	0.08	0.00
Lived with both parents (%)	91	90	91	0.04	0.00
Mother working in 1957 (%)	37	34	36	0.05	0.00
Teachers encouraged college (%)	61	54	59	0.14	0.05
Parents encouraged college (%)	67	67	65	0.00	0.05

[3] Score was mapped from raw Henmon-Nelson scores. See Appendix G – Cor652 for details of the construction

© 2017 American Medical Association. All rights reserved.

Had friends who planned to go to college (%)	44	45	42	-0.01	0.05
Extent to which subject discussed future plans with teachers					
Not at all	25	25	26	0.00	-0.01
Sometimes	63	64	63	-0.03	-0.01
Very much	12	10	11	0.02	0.01
Extent to which subject discussed future plans with parents					
Not at all	3	3	2	0.01	0.01
Sometimes	42	43	45	-0.01	-0.05
Very much	55	54	52	0.01	0.04
Family wealth relative to community					
Considerably below average	1	1	0	0.00	0.00
Somewhat below average	6	8	7	-0.03	-0.01
Average	70	69	69	0.01	0.01
Somewhat above average	21	20	21	0.01	0.00
Considerably above average	3	2	3	0.01	0.00
Financial support from parents for college					
Cannot support (%)	28	33	31	-0.09	-0.05
Can support, with some sacrifice (%)	60	54	56	0.09	0.07
Easily support (%)	13	12	14	0.01	-0.02
Number of ApoE e4 variants	0.30	0.25	0.27	0.10	0.05

eTable 15: Analog to eTable 5 in main text, but restricted to subjects with available ApoE *genotype data. The weighting of after matching values for non-collision sport controls is determined by the matched set composition (eTable 11).*

© 2017 American Medical Association. All rights reserved.

eTable 16. Analog of eTable 6, Restricted to Participants With Available *ApoE* Genotype Data

	Non-Collision Sport Controls	Non-Sport Controls (Before Matching)	Non-Sport Controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957	578.42	547.29	617.31	0.13	-0.17
High school size (N)	148.62	202.33	157.52	-0.41	-0.07
High school rank (quantile)	49.22	45.78	51.67	0.13	-0.09
Parental Income in 1957 ($100)	62.33	62.74	66.50	-0.01	-0.07
Father's Education (years)	10.31	9.70	10.56	0.18	-0.07
Mother's Education (years)	10.80	10.49	11.14	0.11	-0.12
Duncan Socioeconomic Index of father's job in 1957	329.25	342.76	348.32	-0.06	-0.08
Planned future education (years)	2.11	1.73	2.32	0.18	-0.07
IQ[4]	102.68	102.70	104.29	0.00	-0.11
Considered an "outstanding student" by teacher (%)	13	11	14	0.07	-0.03
Participated in band, orchestra, chorus, or musical ensembles (%)	34	26	35	0.17	-0.03
Participated in drama, speech, debate (%)	26	21	30	0.12	-0.09
Participated in school government (%)	27	10	26	0.45	0.03
Participated in school publications (%)	25	13	24	0.30	0.01
Father was a farmer (%)	22	19	21	0.07	0.03
Planned to serve in military (%)	28	27	26	0.03	0.05
Attended Catholic high school (%)	4	13	6	-0.31	-0.05
Lived with both parents (%)	90	90	93	0.01	-0.08
Mother working in 1957 (%)	34	37	33	-0.06	0.01
Teachers encouraged college (%)	54	44	59	0.21	-0.09
Parents encouraged college (%)	67	60	70	0.15	-0.06
Had friends who planned to go to	45	35	51	0.20	-0.14

[4] Score was mapped from raw Henmon-Nelson scores. See Appendix G – Cor652 for details of the construction

© 2017 American Medical Association. All rights reserved.

college (%)					
Extent to which subject discussed future plans with teachers					
Not at all	25	31	23	-0.10	0.03
Sometimes	64	62	66	0.05	-0.03
Very much	10	8	10	0.05	0.00
Extent to which subject discussed future plans with parents					
Not at all	3	2	1	0.01	0.03
Sometimes	43	47	40	-0.07	0.05
Very much	54	51	59	0.06	-0.08
Family wealth relative to community					
Considerably below average	1	0	0	0.01	0.01
Somewhat below average	8	7	7	0.01	0.01
Average	69	68	67	0.01	0.03
Somewhat above average	20	23	24	-0.04	-0.06
Considerably above average	2	2	2	0.01	0.01
Financial support from parents for college					
Cannot support (%)	33	32	28	0.02	0.09
Can support, with some sacrifice (%)	54	55	59	-0.01	-0.07
Easily support (%)	12	13	14	-0.01	-0.02
Number of *ApoE* e4 variants	0.25	0.28	0.24	-0.07	0.00

eTable 16: Analog to eTable 6 in main text, but restricted to subjects with available ApoE *genotype data. The weighting of after matching values for non-sport controls is determined by the matched set composition (eTable 11).*

© 2017 American Medical Association. All rights reserved.

eTable 17. Analog to eTable 7, Restricted to Participants With Available *ApoE* Genotype Data

	Football Players	All Controls (Before matching)	All Controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957 (%)	33	31	32	0.04	0.02
High school rank (%)	3	7	3	-0.16	-0.02
Parental Income in 1957 (%)	3	3	3	0.00	-0.01
Father's Education (years)	1	3	2	-0.11	-0.01
Mother's Education (years)	1	1	1	-0.08	-0.01
Duncan Socioeconomic Index of father's job in 1957	1	1	1	0.00	0.00
Planned future education (years)	8	7	8	0.02	-0.02
Lived with both parents (%)	0	0	0	0.06	0.04
Mother working in 1957 (%)	1	0	0	0.02	0.03
Teachers encouraged college (%)	4	4	5	-0.02	-0.03
Parents encouraged college (%)	3	4	3	-0.05	-0.01
Had friends who planned to go to college (%)	12	12	13	-0.02	-0.02
Extent to which subject discussed future plans with teachers	2	2	2	-0.03	-0.01
Extent to which subject discussed future plans with parents	1	1	1	0.00	0.00
Family wealth relative to community	2	3	2	-0.02	0.01
Financial support from parents for college	12	12	12	-0.01	-0.03

eTable 17: Analog to eTable 7, but restricted to subjects with available ApoE *genotype data. After matching values for all controls are weighted according to the matched set composition (eTable 11).*

© 2017 American Medical Association. All rights reserved.

eTable 18. Analog to eTable 8, Restricted to Participants With Available *ApoE* Genotype Data

	Football Players	Non-sport controls (Before matching)	Non-Sport controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957 (%)	33	32	33	0.02	0.00
High school rank (%)	3	7	3	-0.17	0.02
Parental Income in 1957 (%)	3	3	2	-0.01	0.03
Father's Education (years)	1	3	2	-0.13	0.01
Mother's Education (years)	1	1	1	-0.04	0.00
Duncan Socioeconomic Index of father's job in 1957	1	0	1	0.07	0.02
Planned future education (years)	8	8	8	-0.02	0.01
Lived with both parents (%)	0	0	0	0.05	0.04
Mother working in 1957 (%)	1	0	0	0.05	0.03
Teachers encouraged college (%)	4	4	4	0.01	0.01
Parents encouraged college (%)	3	3	2	-0.04	0.02
Had friends who planned to go to college (%)	12	14	13	-0.05	-0.03
Extent to which subject discussed future plans with teachers	2	2	2	-0.01	-0.01
Extent to which subject discussed future plans with parents	1	1	1	0.01	0.01
Family wealth relative to community	2	3	2	-0.03	0.01
Financial support from parents for college	12	12	13	-0.01	-0.05

eTable 18: Analog to eTable 8 but restricted to subjects with available ApoE *genotype data. After matching values for non-sport controls are weighted according to the matched set composition (eTable 11).*

© 2017 American Medical Association. All rights reserved.

eTable 19. Analog to eTable 9, Restricted to Participants With Available *ApoE* Genotype Data

	Football Players	Non-Collision Sport controls (Before matching)	Non-Collision Sport controls (After Matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957 (%)	33	30	33	0.07	0.01
High school rank (%)	3	7	4	-0.16	-0.03
Parental Income in 1957 (%)	3	2	3	0.04	0.02
Father's Education (years)	1	2	2	-0.07	-0.04
Mother's Education (years)	1	1	1	-0.15	0.02
Duncan Socioeconomic Index of father's job in 1957	1	2	1	-0.09	-0.04
Planned future education (years)	8	5	5	0.12	0.09
Lived with both parents (%)	0	0	0	0.09	0.09
Mother working in 1957 (%)	1	1	0	-0.03	0.04
Teachers encouraged college (%)	4	5	4	-0.07	0.00
Parents encouraged college (%)	3	4	2	-0.06	0.03
Had friends who planned to go to college (%)	12	10	12	0.07	-0.01
Extent to which subject discussed future plans with teachers	2	4	1	-0.06	0.01
Extent to which subject discussed future plans with parents	1	2	2	-0.01	0.02
Family wealth relative to community	2	2	2	0.02	0.02
Financial support from parents for college	12	12	10	-0.02	0.04

eTable 19: Analog to eTable 9, but restricted to subjects with available ApoE *genotype data. After matching values for non-collision sport controls are weighted according to the matched set composition (eTable 11).*

© 2017 American Medical Association. All rights reserved.

eTable 20. Analog to eTable 10, Restricted to Participants With Available *ApoE* Genotype Data

	Non-Collision Sport	Non-Sport Controls (Before Matching)	Non-Sport Controls (After matching)	Standardized Difference (Before Matching)	Standardized Difference (After Matching)
Duncan Socioeconomic Index of job to which subject aspired in 1957 (%)	30	32	33	-0.05	-0.08
High school rank (%)	7	7	5	-0.01	0.06
Parental Income in 1957 (%)	2	3	1	-0.05	0.06
Father's Education (years)	2	3	2	-0.07	0.02
Mother's Education (years)	2	1	1	0.11	0.12
Duncan Socioeconomic Index of father's job in 1957	2	0	1	0.15	0.11
Planned future education (years)	5	8	6	-0.14	-0.04
Lived with both parents (%)	0	0	0	-0.05	-0.11
Mother working in 1957 (%)	1	0	1	0.08	0.04
Teachers encouraged college (%)	5	4	5	0.07	0.02
Parents encouraged college (%)	4	4	4	0.02	0.00
Had friends who planned to go to college (%)	10	14	12	-0.12	-0.07
Extent to which subject discussed future plans with teachers	4	2	3	0.05	0.02
Extent to which subject discussed future plans with parents	2	1	1	0.02	0.03
Family wealth relative to community	2	3	3	-0.05	-0.02
Financial support from parents for college	12	12	10	0.02	0.06

eTable 20: Analog to eTable 10, but restricted to subjects with available ApoE *genotype data. After matching values for non-sport controls are weighted according to the matched set composition (eTable 11).*

© 2017 American Medical Association. All rights reserved.

eTable 21. Analog to Table 2, Restricted to Participants With Available *ApoE* Genotype Data

	Football vs. All Controls	Football vs Non-Sport Controls	Football Players vs. Non-Collision Sport	Other Sport Controls vs. Non-Sport Controls
Combined Cognition Score	-0.01 (-0.11, 0.14)	-0.07 (-0.21, 0.07)	0.07 (-0.08, 0.22)	-0.11 (-0.25, 0.04)
Modified CES-D score	-1.49 (-3.39, 0.41)	-1.22 (-3.20, 0.75)	-1.74 (-3.82, -0.33)	1.27 (-0.92, 3.47)

eTable 21: Analog to Table 2 when restricted to subjects with available ApoE *data. Estimated effect of playing high school football with 97.5% confidence intervals in parentheses. * = significant using the ordered testing procedure. For the composite cognition score, a small effect is 0.16, 0.39, 0.62.. For modified CES-D score, 2.36, 5.91, 9.45*

© 2017 American Medical Association. All rights reserved.

eTable 22. Analog to Table 3, Restricted to Participants With Available *ApoE* Genotype Data

Outcome	Effect (97.5% CI)	Small Effect Cut-off	p-value
Analogs of Primary Outcomes			
Composite Cognition Score (age 72)	-0.00 (-0.13, 0.13)	0.16	0.95
Modified CES-D Score (age 54)	-1.42 (-3.55, 0.71)	2.58	0.70
Modified CES-D Score (age 72)	-0.37 (-2.73, 2.00)	2.72	0.73
Individual Cognitive Test Scores			
Letter Fluency Score (age 65)	0.01 (-0.75, 0.78)	0.86	0.97
Letter Fluency Score (age 72)	0.17 (-0.53, 0.88)	0.83	0.58
Delayed Word Recall Score (age 65)	-0.03 (--0.36, 0.31)	0.39	0.86
Delayed Word Recall Score (age 72)	-0.12 (-0.43, 0.18)	0.32	0.36
Digit Ordering Score (age 65)	-0.15 (-0.66, 0.37)	0.61	0.51
Digit Ordering Score (age 72)	-0.20 (-0.77, 0.37)	0.54	0.44
Similarity Score (age 54)	-0.11 (-0.49, 0.28)	0.56	0.53
Similarity Score (age 65)	-0.03 (-0.37, 0.31)	0.48	0.83
Similarity Score (age 72)	-0.05 (-0.39, 0.30)	0.48	0.77
Immediate Word Recall Score (age 65)	-0.01 (-0.30, 0.27)	0.34	0.91
Immediate Word Recall Score (age 72)	0.00 (-0.26, 0.26)	0.27	0.98
Number Series Score (age 72)	-1.45 (-9.96,7.07)	10.83	0.70
Behavioral and Emotional Outcomes			
Hostility Score (age 54)	0.05 (-0.35, 0.45)	0.47	0.77
Hostility Score (age 65)	-0.16 (-0.49, 0.17)	0.41	0.28
Hostility Score (age 72)	0.10 (-0.29, 0.50)	0.46	0.55
Speilberger Anger Index (age 65)	-0.24 (-1.02, 0.54)	0.95	0.49
Speilberger Anger Index (age 72)	0.37 (-0.47, 1.20)	0.91	0.33
Speilberger Anxiety index (age 65)	-0.35 (-1.45, 0.75)	1.39	0.48
Speilberger Anxiety Index (age 72)	0.10 (-1.37, 1.16)	1.39	0. 85
Heavy Drinking Status (age 54)*‡	--	--	--
Heavy Drinking Status (age 65)*	0.27 (0.07,1.10)	1.5	0.04
Heavy Drinking Status (age 72)*‡	--	--	--

eTable 22: Analog to Table 3 in main text showing estimated effect (CI) of playing football, compared to all controls for our secondary outcomes, restricted to subjects with available ApoE *genotype data. *: Effects on binary outcomes are reported on the odds-ratio scale. The small effect size cutoff corresponding to the Cohen's 0.2 SD threshold is 1.5* [21]*. † = significant at the 5% level after Benjamini-Hochberg correction.‡ = Fitted model was numerically unstable*

© 2017 American Medical Association. All rights reserved.

eTable 23. Estimated Interaction Effect (CI) of Playing High-School Football and *ApoE* e4 Status

	Football only effect	**Football and one *ApoE* e4 Variant joint effect**	**Football and two *ApoE* e4 Variants joint effect**	**Interaction effect of Football and *ApoE* e4 status**
Combined Cognition Score	-0.01 (-0.18, 0.16)	-0.06 (-0.18, 0.30)	0.14 (-0.31, 0.59)	0.08 (0.44)
Modified CES-D score	-1.08 (-3.57, 1.40)	-2.14 (-5.64, 1.35)	-3.56 (-10.19, 3.07)	-1.42 (0.37)

eTable 23: Columns 1-3: Estimated effect of playing high school football, joint effect of playing high school football and having one ApoE *e4 variant, and joint effect of playing high school football and having two* ApoE *e4 variants. 97.5% confidence intervals are reported in parentheses, controlled simultaneously for each outcome measure. Column 4: Estimated interaction effect of playing high school football and* ApoE *e4 status, i.e., the modification of the effect of playing high school football by the presence of an additional* ApoE *e4 variant. Uncorrected p-values reported in parentheses. For the composite cognition score, the cutoffs for small, medium, and large effects are 0.16, 0.41, and 0.65, respectively. For modified CES-D score, the cut-offs are 2.56, 6.41, and 10.25.*

© 2017 American Medical Association. All rights reserved.

eTable 24. Adjusted Differences in Average Cognitive Test and CES-D Scores Over Time for Football Players and Controls

	Letter Fluency (N = 2421)	Delayed Word Recall (N = 2062)	Immediate Word Recall (N = 2086)	Digit Ordering (N = 2051)	Similarity (N = 3209)	CES-D (N = 2971)
Age (Control)						
65 – 54	NA	NA	NA	NA	-1.35 (<0.001†)	-2.52 (<0.001†)
72 – 54	NA	NA	NA	NA	-1.58 (<0.001†)	-0.19 (0.60)
72 – 65	-0.39 (0.001†)	-0.55 (<0.001†)	-0.47 (<0.001†)	-0.08 (0.44)	-0.22 (<0.001†)	2.34 (<0.001†)
Age (Football)						
65 – 54	NA	NA	NA	NA	-1.25 (<0.001†)	-2.76 (<0.001†)
72 – 54	NA	NA	NA	NA	-1.69 (<0.001†)	0.36 (0.48)
72 – 65	-0.24 (0.16)	-0.46 (<0.001†)	-0.49 (<0.001†)	-0.17 (0.26)	-0.44 (<0.001†)	3.13 (<0.001†)
Group Diff (FB – Control)						
65 – 54	NA	NA	NA	NA	0.10 (0.32)	-0.24 (0.68)
72 – 54	NA	NA	NA	NA	-0.11 (0.32)	0.55 (0.38)
72 – 65	0.15 (0.45)	0.09 (0.43)	-0.01 (0.88)	-0.09 (0.62)	-0.21 (0.06)	0.79 (0.21)
Effect of Football						
Age 54	NA	NA	NA	NA	-0.10 (0.28)	-1.18 (0.05)
Age 65	-0.30 (0.14)	-0.09 (0.31)	0.03 (0.66)	0.03 (0.81)	0.01 (0.95)	-1.41 (0.02)
Age 72	-0.15 (0.44)	-0.01 (0.95)	0.02 (0.81)	-0.06 (0.72)	-0.21 (0.05)	-0.63 (0.34)

Table 1: Effect of aging from 54 to 65 to 72 among all controls and football players on each cognitive test and CES-D score. Also shown are between-group differences in aging and the overall effect of playing football at each age. † = significant at 0.05 level after BH correction

© 2017 American Medical Association. All rights reserved.

eReferences

1. Lezak MD, Howieson DB, Bigler ED, Tranel D. *Neuropsychological Assessment (5th Ed.)*.; 2012.
2. Belanger HG, Curtiss G, Demery J a, Lebowitz BK, Vanderploeg RD. Factors moderating neuropsychological outcomes following mild traumatic brain injury: a meta-analysis. *J Int Neuropsychol Soc*. 2005;11(3):215-227. doi:10.1017/S1355617705050277.
3. Wilde EA, Whiteneck GG, Bogner J, et al. Recommendations for the use of common outcome measures in traumatic brain injury research. *Arch Phys Med Rehabil*. 2010;91(11). doi:10.1016/j.apmr.2010.06.033.
4. Belanger HG, Vanderploeg RD. The neuropsychological impact of sports-related concussion: A meta-analysis. *J Int Neuropsychol Soc*. 2005;11:345-357. doi:10.1017/S1355617705050411.
5. Guskiewicz KM, Marshall SW, Bailes J, et al. Association between recurrent concussion and late-life cognitive impairment in retired professional football players. *Neurosurgery*. 2005;57(4):719-726. doi:10.1227/01.NEU.0000175725.75780.DD.
6. Radloff L. The CES-D Scale: A self-report depression scale for use in general populations. *Appl Psychol Meas*. 1977;1:385-400.
7. Spielberger CD, Gorsuch R, Lushene RE, Vagg P, Jacobs GA. *Manual for the State-Trait Anxiety Inventory (Form Y1)*. CA: Consulting Psychologists Press; 1983.
8. Spielberger CD. *State-Trait Anger Expression Inventory*. Lutz, FL: Psychological Assessment Resources; 1999.
9. Rosenbaum PR. A characterization of optimal designs for observational studies. *J R Stat Soc Ser B*. 1991;53(3):597-610.
10. Rosenbaum PR. *Design of Observational Studies*. New York: Springer; 2010.
11. Rosenbaum PR, Rubin DB. Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. *Am Stat*. 1985;39(1):33-38. doi:10.1080/00031305.1985.10479383.
12. Gu XS, Rosenbaum PR. Comparison of Multivariate Matching Methods☐: Structures, Distances, and Algorithms. *J Comput Graph Stat*. 1993;2(4):405-420. doi:10.1080/10618600.1993.10474623.
13. Silber JH, Rosenbaum PR, Trudeau ME, et al. Multivariate matching and bias reduction in the surgical outcomes study. *Med Care*. 2001;39(10):1048-1064. doi:10.1097/00005650-200110000-00003.
14. Cochran WG, Rubin DB. Controlling Bias in Observational Studies☐: A Review. *Sankhya Indian J Stat Ser A1*. 1973;35(4):417-446.
15. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc*. 1984;79(387):516-524.
16. Hofmann H, Kafadar K, Wickham H. *Letter-Value Plots: Boxplots for Large Data*.; 2011.
17. Tukey JW. *Exploratory Data Analysis*. 1st ed. Addison-Wesley; 1977.
18. Hoaglin DC. Letter values: A set of selected order statistics. In: *Understanding Robust and Exploratory Data Analysis*. New York: Wiley; :33-57.
19. Rosenbaum PR. Testing Hypothesis in Order. *Biometrika*. 2008;95(1):248-252. doi:10.1093/biomet/asm085.

© 2017 American Medical Association. All rights reserved.

20. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Erlbaum; 1988.
21. Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Commun Stat - Simul Comput*. 2010;39(4):860-864. doi:10.1080/03610911003650383.

© 2017 American Medical Association. All rights reserved.