AUS DER

KLINIK UND POLIKLINIK FÜR KINDER- UND JUGENDSPSYCHIATRIE, PSYCHOSOMATIK UND PSYCHOTHERAPIE

KLINIKUM DER LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

VORSTAND: PROF. DR. GERD SCHULTE-KÖRNE

SUBCORTICAL GRAY MATTER ALTERATIONS IN FORMER PROFESSIONAL AMERICAN FOOTBALL PLAYERS

DISSERTATION

ZUM ERWERB DES DOKTORGRADES DER MEDIZIN
AN DER MEDIZINISCHEN FAKULTÄT DER
LUDWIG-MAXIMILIANS-UNIVERSITÄT ZU
MÜNCHEN
VORGELEGT VON

VIVIAN SCHULTZ

AUS
BAD OLDESLOE

2022

MIT GENEHMIGUNG DER MEDIZINISCHEN FAKULTÄT

DER UNIVERSITÄT MÜNCHEN

Erster Gutachter:	Prof. Dr. Inga Katharina Koerte
Zweiter Gutachter:	Prof. Dr. Nikolaos Koutsouleris
Dritter Gutachter:	Prof. Dr. Peter zu Eulenburg
Weitere Gutachter:	Prof. Dr. Thomas Liebig
Mitbetreuung durch den promovierten Mitarbeiter:	
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	20.01.2022

AFFIDAVIT

Eidesstattliche Versicherung

Schultz, Vivian

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Thema

“Subcortical Gray Matter Alterations in Former Professional American Football Players”

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 14.03.22	Vivian Schultz
---	---
Ort, Datum	Unterschrift Doktorandin

TABLE OF CONTENTS

Affidavit - 3 -

Table of Contents - 4 -

List of Figures - 6 -

List of Tables - 7 -

Abbreviations - 8 -

Publication Record of Presented Work - 9 -

English Summary - 10 -

Deutsche Zusammenfassung - 12 -

1. Introduction - 15 -

1.1 Chronic Traumatic Encephalopathy - 15 -

1.1.1 Overview - 15 -

1.1.2 Pathology and Diagnosis - 15 -

1.1.3 Possible Neurodegenerative Cascade - 16 -

1.1.4 Clinical Presentation - 16 -

1.1.5 Approaches for Diagnosis In Vivo - 17 -

1.1.6 Subcortical Gray Matter In RHI and CTE - 18 -

1.1.7 Risk Factors - 18 -

1.1.8 Treatment and Prevention - 20 -

2. Methodological Considerations - 21 -

2.1 Magnetic Resonance Imaging - 21 -

2.2 Quantification of Brain Regional Volumes - 21 -

2.3 Freesurfer - 22 -

2.3.1 Overview - 22 -

2.3.2 Pre-Processing - 22 -

2.3.3 Label Construction and Assignment - 22 -

2.3.4 Possible Limitations - 23 -

2.4 Neurobehavioral Assessment - 24 -

3. Research Questions and Hypotheses - 28 -

4. Publication A - 30 -

4.1 Background - 30 -

4.2 Methods - 30 -

4.3 Results - 30 -

4.4 Discussion - 31 -

4.5 Contribution - 32 -

5. Publication B - 33 -

5.1 Background - 33 -

5.2 Methods - 33 -
5.3 Results - 33 -
5.4 Discussion - 33 -
5.5 Contribution - 35 -
References **- 36 -**

LIST OF FIGURES

Fig. 1 – Neuropathologic stages of CTE

Fig. 2 – FreeSurfer processing stream

LIST OF TABLES

Table 1 – Neuropsychological testing battery

Table 2 – Mean differences of volume between NFL players and controls adapted from Publication A, 151

ABBREVIATIONS

AFE	Age at First Exposure
BQSS	Boston Qualitative Scoring System
CSF	Cerebro-Spinal Fluid
CTE	Chronic Traumatic Encephalopathy
DSM-V	Diagnostic and Statistical Manual of Mental Disorders
FID	Free-Induction Decay
ICC	Intraclass Correlation Coefficient
MRI	Magnetic Resonance Imaging
MRT	Magnetresonanztomographie
NFL	National Football League
NINDS	National Institute on Neurological Disorders and Stroke
NIBIB	National Institute of Biomedical Imaging and Bioengineering
NMR	Nuclear Magnetic Resonance
P-Tau	Hyperphosphorylated Tau
RHI	Repetitive Head impacts
ROI	Region of Interest
SAS	Statistical Analysis System
TES	Traumatic Encephalopathy Syndrome

PUBLICATION RECORD OF PRESENTED WORK

This dissertation is based on the following two publications:

1. Guenette, J.P., Stern, R.A., Tripodis, Y., Chua, A.S., **Schultz, V.**, Sydnor, V.J., Somes, N., Karmacharya, S., Lepage, C., Wrobel, P., Alosco, M.L., Martin, B.M., Chaisson, C.E., Coleman, M.J., Lin, A.P., Pasternak, O., Makris, N., Shenton, M.E., Koerte, I.K.:
Automated Versus Manual Segmentation of Brain Region Volumes in Former Football Players, Neuroimage: Clinical, 2018

- Impact factor 2018: 3.94

2. **Schultz, V.**, Stern, R.A., Tripodis, Y., Stamm, J.S., Wrobel, P., LePage, C., Weir, I., Baugh, C., Fritts, N.G., Martin, B.M., Chaisson, C., Coleman, M.J., Lin, A.P., Pasternak, O., Shenton, M.E.*, Koerte, I.K.*:
Age at First Exposure to Repetitive Head Impacts Is Associated with Smaller Thalamic Volumes in Former Professional American Football Players, Journal of Neurotrauma, 2018.

- Impact factor 2018: 3.75

ENGLISH SUMMARY

Chronic traumatic encephalopathy (CTE) is a serious neurodegenerative disorder which has been associated with a history of exposure to repetitive head impacts (RHI). Next-of-kin information clinically characterized CTE as a progressive impairment in mood and behavior along with a decline in cognitive and motor function. Contact-sport athletes are at increased risk. To date, CTE can only be diagnosed *postmortem*. Histologically, it is characterized by deposition of hyperphosphorylated tau in the depth of the sulci. Among the macroscopic pathologic findings is subcortical volume loss, including atrophy of the thalamus, a central subcortical structure involved in multiple cortex-controlled pathways. Risk and modulating factors of CTE other than exposure to RHI are largely unknown. However, recent data suggest that in particular the developing brain has a high vulnerability to RHI. In fact, younger age at first exposure (AFE) to contact-sports correlated not only with more neuropsychiatric symptoms later in life and microstructural white matter alterations, but also with earlier symptom onset in CTE in *postmortem* studies. Therefore, AFE to contact-sports should be evaluated as a possible risk factor for CTE, especially considering the large number of youth athletes participating in contact-sports today.

Neuroimaging has enabled researchers to gain insight into sequelae of RHI *in vivo*. In particular, analyses of global and regional brain volumes have shed light on inter-individual volumetric differences and age-dependent volumetric changes in contrast to pathologic forms of brain atrophy. Especially advanced magnetic resonance imaging (MRI) techniques have made it possible to study and quantify even subtle of the aforementioned changes. Before automated segmentation software was available, brain global and regional volumes had to be derived manually. In fact, until today manual tracing is considered to be the gold standard for calculating brain volumes in research studies. However, manually segmenting a brain based on an MRI scan is time consuming. Given the large datasets studies to date often involve, automatic segmentation programs pose promising and more time-efficient alternatives to manual tracing. A software routinely applied in this context to automatically segment the brain is FreeSurfer. While time efficiency is a major advantage of automatic segmentation software, the same or at least greatly similar anatomical accuracy and especially reliability compared with manual tracing is demanded. Surprisingly, only a small number of studies have tested FreeSurfer's reliability in comparison to manual tracing. Results from these studies remain mixed. Especially in subjects with significant volumetric divergence from the norm or with anatomical anomalies, FreeSurfer has shown weaknesses and significantly differing results from manual tracing. This is of particular note, as volumetric analyses play a major role in the research of RHI and possible CTE which is associated with characteristic cortical and subcortical atrophy patterns and anatomical variances.

In light of the above, the goal of the work presented in this thesis is to explore the accuracy of FreeSurfer compared with manual tracing and to investigate the association between thalamic volumes and AFE to American Football in a sample of symptomatic former professional American Football players at risk of CTE.

Publication A: The purpose of this study was to compare volumetric results by FreeSurfer 5.3 with manual tracing in a sample at risk of CTE. The results were compared in terms of anatomical accuracy, distinction between players and a control group, and associations with clinical measures.

A unique sample of eighty-six symptomatic former professional football players with a minimum of two years of participation in the National Football League (NFL) as well as twenty-two age-matched healthy controls without a history of contact-sports or brain trauma were included. Volumes of eleven brain regions often affected in previous studies of RHI were derived: bilateral corpus callosum, cingulate gyrus, amygdala, hippocampus, amygdala-hippocampal complex, and lateral ventricles. Volumes were obtained by both, automated segmentation and manual correction of the FreeSurfer labels.

Inter-observer-reliability showed good to excellent results. The correlations of the lateral ventricles and the corpus callosum were excellent between automatically and manually derived volumes. The results for the amygdala, hippocampus, and cingulate gyrus, however, were more divergent. The differences were significant in terms of distinction between groups and correlations with clinical measures. Differences between football players and controls were found in 8 of 11 manually-corrected regions versus only 4 of 11 regions without manual edit. Among the football players, the correlation between brain regional volumes and clinical measures led to different results for manual correction versus automated segmentation. The results suggest that amygdala, hippocampus, and cingulate gyrus should be manually edited prior to data analysis in a sample at risk of CTE. Better results were obtained if amygdala and hippocampus were combined as amygdala-hippocampal complex which should be considered in studies of RHI and where it is in line with the research question.

Publication B: The goal of this project was to explore the association between AFE to American Football and thalamic volumes in a sample of former professional football players at risk of CTE. We also investigated the association between total years of playing football and thalamic volumes as well as the association with clinical measures.

Thalamic volumes from the same eighty-six retired professional football players from publication A were derived with FreeSurfer 5.3. This study showed an association between thalamic volumes and total duration of playing football indicating that the longer athletes played football, the smaller their thalamic volumes. Further, this is the first study showing an association between subcortical gray matter volume and AFE to contact-sports. Right thalamic volume was associated with AFE, even when adjusting for total years of play. Thus, the younger athletes were when they started playing, the smaller their right thalamic volume. Importantly, the effect of AFE on thalamic volume was almost twice the one of total years of play. This result indicates that RHI may have an especially strong impact on the developing brain potentially due to interference with critical periods of brain development and maturation. More mood and behavioral symptoms as well as worse performance in visual memory was linked to thalamic volume in this cohort at risk of CTE. The findings suggest that symptoms associated with CTE may be related to decreased thalamic volume. However, without *postmortem* analysis, it is not possible to determine whether or not the findings of this study are specifically linked to CTE.

DEUTSCHE ZUSAMMENFASSUNG

Die Chronisch Traumatische Enzephalopathie (CTE) ist eine schwere neurodegenerative Erkrankung, die mit wiederholten Kopferschütterungen assoziiert ist. Klinisch zeichnet sie sich durch progressive kognitive und motorische Einschränkungen sowie Stimmungs- und Verhaltensauffälligkeiten aus. Kontaktsportathleten sind eine der größten bekannten Risikogruppen. Bis heute kann die Diagnose einer CTE nur *postmortal* erfolgen. Pathognomonisch sind Läsionen mit hyperphosphoryliertem Tau in den Tiefen der Sulci. Zum typischen makroskopisch-pathologischen Erscheinungsbild zählen subcorticale Volumenreduktionen, unter anderem des Thalamus, einer zentralen subcortikalen Struktur, die in zahlreiche cortical gesteuerte Prozesse involviert ist. Ergebnisse einer Studie unter professionellen Kampfsportathleten haben zudem eine Verbindung zwischen Reduktion des Thalamusvolumens und Exposition gegenüber Kopferschütterungen *in vivo* gezeigt. Neben Kopferschütterungen sind weitere, die Erkrankung beeinflussende und Risikofaktoren weitestgehend unbekannt. Neuere Daten zeigen, dass insbesondere das sich noch in der Entwicklung befindende Gehirn eine erhöhte Vulnerabilität gegenüber Kopferschütterungen aufweist. Daher sollte das Startalter von Kontaktsportathleten als möglicher Risikofaktor für CTE untersucht werden, vor allem vor dem Hintergrund der großen Anzahl Kinder und Jugendlicher, die aktuell Kontaktsportarten betreiben.

Der Einsatz von *Neuroimaging* hat Wissenschaftlern Einblicke in Folgeschäden von Kopferschütterungen *in vivo* ermöglicht. Insbesondere Volumenanalysen erlauben die Unterscheidung von interindividuellen Normvarianten und altersassoziierten Veränderungen vom Übergang ins Pathologische. Besonders moderne Techniken der Magnetresonanztomographie (MRT) machen es möglich, diese z.T. kleinsten Veränderungen zu quantifizieren. Bevor es automatisierte Segmentierungssoftware gab, mussten Hirnvolumina manuell extrahiert werden. Tatsächlich ist die manuelle Volumenextraktion nach wie vor der Goldstandard in wissenschaftlichen MRT-basierten Volumenanalysen. Jedoch ist die manuelle Segmentierung sehr zeitaufwendig. Insbesondere aufgrund der großen Datensätze, die wissenschaftliche Studien heutzutage in der Regel umfassen, stellen automatisierte Segmentierungssoftwares eine vielversprechende und zeiteffizientere Alternative dar. Eine Software, die in diesem Zusammenhang routinemäßig zum Einsatz kommt, ist FreeSurfer. Während jedoch Zeiteffizienz ein großer Vorteil der automatisierten Segmentierung ist, so ist doch die anatomische Genauigkeit und die Reliabilität Voraussetzung für deren Einsatz. Überraschenderweise haben bislang nur wenige Studien FreeSurfers Reliabilität im Vergleich zur manuellen Segmentierung untersucht und sind hierbei zu unterschiedlichen Ergebnissen gelangt. Vor allem bei Probanden mit großen volumetrischen Abweichungen von der Norm oder strukturellen anatomischen Anomalien hat FreeSurfer Schwächen und signifikant abweichende Ergebnisse von der manuellen Segmentierung gezeigt. Dies gilt es vor allem deshalb zu beachten, da Volumenanalysen insbesondere bei der Erforschung der Folgen von Kopferschütterungen und potentiellen Fällen von CTE eine wichtige Rolle spielen, die typischerweise charakteristische corticale und subcorticale Atrophiemuster und anatomische Anomalien aufweisen.

Aus dem oben Genannten ergeben sich die Ziele der im Folgenden vorgestellten Studien. Es soll zunächst die Genauigkeit von FreeSurfer in einer Kohorte symptomatischer, ehemaliger professioneller American Football

Spieler untersucht werden, die eine Risikopopulation für CTE darstellen. Darauffolgend wird der Zusammenhang zwischen Thalamusvolumen und Startalter im Kontaktsport in derselben Kohorte analysiert.

Publikation A: Das Ziel dieser Studie war es FreeSurfer 5.3 mit manueller Nachbearbeitung der Segmentierung bei Volumenanalysen in einer Risikopopulation für CTE zu vergleichen. Die Ergebnisse wurden hinsichtlich der anatomischen Genauigkeit, der Unterscheidung zwischen American Football Spielern und Kontrollprobanden sowie der Korrelation mit klinischen Parametern verglichen.
Eine weltweit einzigartige Kohorte aus 86 symptomatischen, ehemaligen American Football Spielern mit mindestens zwei Jahren Erfahrung in der NFL sowie 22 Kontrollprobanden ohne Schädel-Hirn-Trauma oder Teilnahme an Kontaktsportarten in der Anamnese wurden eingeschlossen. Die jeweils bilateralen Volumina von elf Hirnstrukturen wurden extrahiert, ausgewählt anhand vorheriger Studien zu Kopferschütterungen: Corpus callosum, Gyrus cinguli, Amygdala, Hippocampus, Amygdala-Hippocampus-Komplex und die Seitenventrikel. Die Volumina wurden jeweils einmal voll automatisch mithilfe von FreeSurfer 5.3 extrahiert und ein zweites Mal einer manuellen Nachbearbeitung unterzogen.
Die Übereinstimmungsrate bei der manuellen Nachbearbeitung war gut bis ausgezeichnet in dieser Studie. Die Übereinstimmung von automatischer Segmentierung und manueller Nachbearbeitung war hervorragend für Corpus callosum und Seitenventrikel, nicht jedoch für Amygdala, Hippocampus und Gyrus cinguli. Die Unterschiede waren signifikant hinsichtlich der folgenden Gruppenanalysen sowie Korrelationen mit klinischen Parametern. Es wurden Gruppenunterschiede in 8 von 11 manuell nachbearbeiteten Hirnstrukturen gefunden, jedoch nur in 4 von 11 automatisch segmentierten Hirnstrukturen. Die Korrelationen zwischen Volumina und klinischen Parametern innerhalb der Athletengruppe zeigte für manuelle und automatisch segmentierte Volumina verschiedene Ergebnisse. Die Ergebnisse sprechen dafür, dass Amygdala, Hippocampus und Gyrus cinguli in Risikopopulationen für CTE vor weiteren Analysen einer manuellen Nachbearbeitung unterzogen werden sollten. Zudem wurde eine bessere Übereinstimmung erzielt, wenn Amygdala und Hippocampus als Amygdala-Hippocampus-Komplex zusammengefasst wurden, was in zukünftigen Studien zu Kopferschütterungen und wo es die wissenschaftliche Fragestellung erlaubt, berücksichtigt werden sollte.

Publikation B: Das Ziel dieser Studie war es, den Zusammenhang zwischen dem Startalter des Kontaktsports und dem Thalamusvolumen zu untersuchen. Zudem wurden die Thalamusvolumina mit der Gesamtspielzeit in Jahren sowie den Leistungen in diversen neuropsychologischen Tests in Zusammenhang gesetzt.
Es wurden die Thalamusvolumina derselben 86 ehemaligen professionellen American Football Spieler aus Publikation A mithilfe von FreeSurfer 5.3 extrahiert. Es zeigte sich ein Zusammenhang zwischen der Gesamtspielzeit in Jahren und den Thalamusvolumina, insofern als dass je länger die Athleten insgesamt American Football gespielt hatten, desto kleiner deren Volumina. Dies ist darüber hinaus die erste Studie, die einen Zusammenhang zwischen Startalter des Kontaktsports und subcorticalem Volumen der grauen Substanz gezeigt hat. Das Volumen des rechten Thalamus korrelierte mit dem Alter, in dem erstmals American Football gespielt worden war, auch dann, wenn für die Gesamtspielzeit in Jahren statistisch kontrolliert wurde. Je früher

die Athleten American Football gespielt hatten, desto kleiner ihr rechter Thalamus. Dabei ist hervorzuheben, dass der Effekt des Alters zu Beginn der Exposition auf das Thalamusvolumen fast doppelt so stark war wie der der Gesamtspielzeit in Jahren. Die Ergebnisse können so interpretiert werden, dass wiederholte Kopferschütterungen eine besonders starke Auswirkung auf das Gehirn von Kindern oder Jugendlichen haben könnten, möglicherweise aufgrund von Wechselwirkungen mit cerebralen Reifungsprozessen. Zudem korrelierten in dieser Studie stärker ausgeprägte Stimmungs- und Verhaltensauffälligkeiten und Defizite des visuellen Gedächtnisses mit dem Thalamusvolumen. Die Ergebnisse sprechen dafür, dass die für CTE typischen Symptome einen Zusammenhang mit dem Thalamusvolumen zeigen. Es bleibt jedoch zu klären, ob die Ergebnisse tatsächlich mit CTE assoziiert sind, da dies ohne *postmortem* Untersuchung zum aktuellen Zeitpunkt nicht möglich ist.

1. INTRODUCTION

1.1 Chronic Traumatic Encephalopathy

1.1.1 Overview

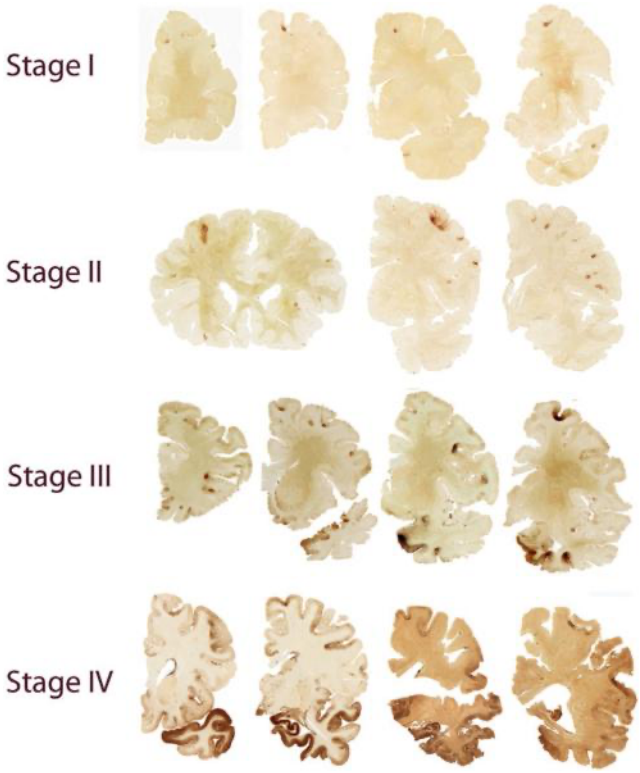
CTE is a severe neurodegenerative disorder linked to a history with repetitive head trauma.[1-3] CTE was clinically first described 1928 by US pathologist Harrison Martland who observed and described neurologic sequelae in boxers.[4] Later on, the condition was referred to as "punch drunk syndrome" and "dementia pugilistica".[5] The name "chronic traumatic encephalopathy" was first introduced in 1949 when it became clear that the symptoms formerly described by Martland were not limited to consequences of boxing.[6] In fact, until today, CTE has most often been described in former contact-sport athletes, such as American Football players.[7-10] Further, CTE has been found following other forms of RHI such as blast events in military service, physical abuse, autistic patients with head banging behavior, and epileptic seizures.[2,11-15]

CTE has severe consequences for patients and relatives. [16] While CTE until now has most often been diagnosed in former professional contact-sport athletes, more recent evidence has shown that also amateur contact-sport athletes may be affected.[17] Therefore, in spite of currently unknown incidence and prevalence of CTE [3], possibly millions are at risk. Further, considering these potentially large patient numbers, CTE may also pose a serious economic burden when considering the costs for dementias other than Alzheimer's Disease in the USA at 73 billion US Dollars annually.[18]

1.1.2 Pathology and diagnosis

To date, CTE can only be diagnosed at *postmortem* neuropathological examination.[2,19] CTE is a tauopathy and shows a distinctive neuropathological pattern which was addressed in the first National Institute on Neurological Disorders and Stroke (NINDS) / National Institute of Biomedical Imaging and Bioengineering (NIBIB) consensus meeting to establish diagnostic criteria for CTE.[20] Based on these criteria, required pathognomonic finding is a lesion related to aggregates of hyperphosphorylated tau (p-tau) in neurons and astrocytes close to small vessels at the depths of the sulci.[20] McKee et al. described four stages of CTE.[19] In early stages, CTE presents as a cortical disease. The perivascular lesions usually first appear in the frontal and temporal lobes, less often in the parietal lobe. In later stages, cortical lesions become more frequent and may spread over the entire cortex. Further, at this stage, also subcortical structures become affected.[19] P-tau lesions may then be present in the mid-temporal lobe, especially in the entorhinal cortex, amygdala, and hippocampus.[3] In later stages, there are additional non-p-tau related pathologies. Common macroscopic features observed include enlargement of the third ventricle, cavum septum pellucidum, as well as mammillary body and thalamic volume loss.[3,21] Finally, global brain atrophy has been described in later stages.[3]

Fig. 1 The four neuropathologically defined stages of CTE from McKee et al.[22].



1.1.3 Possible neurodegenerative cascade

Tau's function in the healthy is well known, namely stabilizing microtubules in the axons of neurons.[23] The mechanism of tau becoming pathologic, however, is not well understood. Tau protein has two conformations, physiologic trans- and pathologic cis-p-tau.[24] In the healthy, the prolyl-isomerase-Pin1 protects against tauopathy by converting p-tau from cis- to trans-isomer.[24] Recent studies using antibodies specifically designed to either bind to cis- or trans-conformation of p-tau suggest that the cis-conformation of tau protein may initiate the neurodegenerative cascade in CTE.[24,25] Findings from brain trauma in mice have shown that hours after impact or neuronal stress, neurons particularly produce cis-p-tau.[25] Over the course of time, cis-p-tau spreads over the cortex, and later to other regions of the brain, inducing apoptosis, and thereby eventually leading to atrophy and neurodegeneration.[25]

1.1.4 Clinical presentation

While the definite diagnosis of CTE can only be made after death, conclusions about clinical signs and course of the disease are drawn from retrospective analyses and interviews with relatives.

CTE is a severe disorder with increasing impairment in everyday life. Four domains of symptoms have been reported: CTE may present with progressive mood, behavioral, cognitive, and motor decline.[21,26,27] Symptom onset in CTE seems to become apparent 8 to 10 years after the end of exposure to RHI often after a symptom-free phase.[2,3] Clinical presentation in CTE varies and shows a large overlap with other psychiatric and neurodegenerative disorders.[28] Stern et al. described two stages of clinical presentation, an early and a late

onset.[16] The early manifestation of CTE (mean age of symptom onset: 35 years) first presents with mood and behavioral symptoms, such as verbal and physical violence, loss of control, impulsivity, and paranoia. Over the course of the disease, these "early-onset" patients additionally develop cognitive symptoms. Patients with a late onset of CTE (mean age of symptom onset: 60 years) first present with cognitive symptoms, such as deficits in memory and executive functioning as well as attention problems. These "late-onset" patients may develop mood and behavioral symptoms with progress of the disease.[16] Approximately 45% of all patients suffering from CTE have dementia.[16] Further, patients may develop impaired motor function, such as parkinsonism, dysarthria, dysphagia, and problems in coordination.[3]

1.1.5 Approaches for diagnosis in vivo

One of the biggest obstacles on the way to a better understanding of CTE is the current lack of a diagnosis *in vivo*. So far, four different clinical approaches have been proposed to establish diagnostic criteria for CTE in the living.[9,29-31] Jordan et al. were the first to introduce possible diagnostic guidelines based on symptom presentation.[29] They differentiate four types of clinical diagnoses: definite, probable, possible, and improbable CTE. The aim of these diagnostic criteria is to predict underlying CTE pathology with the diagnosis "definite CTE" still being limited for *postmortem*.[29] Victoroff et al. and Montenigro et al. focus on grouping and ranking sets of symptoms.[9,31] Montenigro et al. introduced the term "traumatic encephalopathy syndrome (TES)" for clinical diagnosis, emphasizing the difference between the definite neuropathological diagnosis of CTE and the introduced clinical guidelines for TES.[9] The latest clinical diagnostic criteria have been introduced by Reams and colleagues.[30] In analogy with previous diagnostic criteria, there is no category of a definite diagnosis of CTE which still is being reserved for neuropathologic examination.[30]

Although these are important first steps, the use of these guidelines is not established in clinical settings and their specificity remains to be elucidated.[21] These symptom-based approaches show a great overlap with other neurodegenerative diseases[21] and may postpone diagnoses to more advanced stages of the disease where patients are more likely to present with higher symptom loads. The use of neuroimaging techniques to the contrary, provides a reproducible and objective method to investigate brain structural sequelae in those at risk of CTE.[32,33] MRI has been shown sensitive to even subtle alterations after RHI[34,35], forgoes without radiation, and is widely available. Identifying imaging markers of CTE using MRI would vastly promote our understanding of the disease. First, an imaging marker would allow a diagnosis *in vivo* and in early stages.[32] Second, earlier and definite diagnosis in the living would lead to more targeted clinical trials by including only those who suffer from CTE. Third, intervention and treatment options could be elaborated and tested more precisely. Subsequently, this would allow identification of possible risk factors, prediction of the course of the disease, evaluation of the efficacy of treatment and interventions, and, ultimately, protection of those who are at risk.

1.1.6 Subcortical gray matter in RHI and CTE

Atrophy of subcortical gray matter structures has been reported in neuropathologically confirmed cases of CTE.[19] Depending on the stage of CTE, this may include the amygdala, hippocampus, cingulate gyrus, and the thalamus.[19] The *postmortem* evidence of focal atrophy highlights the importance of volumetric analyses *in vivo*. In fact, brain regional volume reduction has been reported in living contact-sport athletes who were exposed to RHI. For example, volume reduction of the hippocampus has been reported among boxers, mixed martial arts fighters as well as American Football players[36-38] whereas volume reduction of the amygdala has been observed in a previous study among professional fighters[36,37]. In another project which is not part of this thesis, I analyzed limbic system structure volumes in a sample of former professional football players at risk of CTE as well as age-matched healthy controls.[39] In fact, the same football players who were introduced in the two publications of this dissertation were included. The volumes of hippocampus, amygdala, and cingulate gyrus were compared between football players and controls who had no prior brain trauma or contact-sport experience. Results from this study showed reduced volumes of amygdala, hippocampus, and cingulate gyrus in former NFL players compared with controls.[39] Moreover, we showed that lower volumes of hippocampus and cingulate gyrus were associated with worse executive function and visual memory among the players' group.[39] The findings are in line with the literature and suggest that the integrity of limbic system structures may be affected through potential neurodegenerative processes associated with RHI and that alterations of limbic system structures are associated with symptoms typical of CTE.

In spite of these results, and while thalamic atrophy has been reported in neuropathologically confirmed cases of CTE, analyses of the thalamus in living athletes are still sparse. The thalamus is a central subcortical structure which is involved in multiple cortex-controlled pathways and has been described as the "gateway to consciousness"[40]. Anatomically, the thalamus (thalamus dorsalis) consists of several subregions or nuclear complexes.[41] Thalamic nuclei can be classified in different ways. One possibility is to divide them according to functional classes.[42] Five different classes have been described: intralaminar and reticular nuclei for nociception and arousal, effector nuclei associated with motor function, sensory nuclei involved in all major thalamic domains, limbic nuclei for mood/behavior, and associative nuclei involved in cognition.[42] To date, only Bernick and colleagues investigated the association between thalamic volume and exposure to RHI in the living.[36,37,43] They included a sample of professional boxers and mixed martial arts fighters and found a reduction of thalamic volumes with every fight and with total years of fighting.[36,43] Whether or not thalamic volumes show an association with RHI in American Football, where players are equipped with helmets and head accelerations are reported to be lower than in fighting[44,45] was analyzed and discussed in publication B.

1.1.7 Risk factors

CTE has most often been reported in former professional contact-sport athletes such as American Football players who were exposed to RHI.[7-9] To date, a history of exposure to RHI is presumed to be one of the most important risk factors for developing CTE.[1,3,32,46] RHI refer to multiple, periodically recurring impacts to the head. An impact is a hit to the head that may or may not result in immediate symptoms.[35] Impacts that lead to

acute symptoms such as headache, nausea, dizziness, insomnia, postural instability, phono- and photophobia, blurred vision, fatigue, concentration, and memory problems are referred to as concussive impacts.[47] Impacts that do not lead to acute symptoms are often called subconcussive impacts. Subconcussive impacts are far more common than concussive impacts.[35] There is evidence that especially the repetitive character of head impacts may have adverse effects on the brain.[27,35] However, not everyone who was exposed to RHI develops CTE suggesting that exposure to RHI alone is not sufficient for the pathogenesis of CTE.[3,48,49] Additional risk factors likely influence the susceptibility to RHI and support the development of CTE. Identifying these factors is one of the key questions in CTE research.[8] In the following, two possible risk factors will be discussed.

Total years of play

It has been shown that the total impact burden plays an important role in the susceptibility for developing CTE.[26,50,51] Importantly, both concussive as well as subconcussive head impacts (cumulatively RHI) need to be taken into consideration when considering the total impact burden based on a paper that reported that 16% of neuropathologically confirmed cases of CTE had no history of a concussion.[52] It is methodically challenging to quantify the total amount of RHI an athlete has been exposed to.[53] There are different ways to account for exposure to RHI in prospective study designs. Contact-sport athletes may be equipped with helmet accelerometers or head hits could be counted by video surveillance of trainings and matches. The vast majority of CTE studies until today, however, have retrospective study designs which rely on next-of-kin information or self-report[26,54] making the report of each hit to the head practically impossible. Among contact-sport athletes a possibility to account for the two aforementioned issues is to use the total years the athlete has actively participated in his field of sport as a measure of exposure. In the context of retrospective study designs, the use of total years as a variable poses the advantage of feasibility by being a realistically ascertainable number. In two previous *postmortem* studies, McKee and colleagues not only showed that longer duration of playing football increased the odds of CTE but also that it was associated with more severe neuropathologic stages.[19,51] Importantly, an *in vivo* study among collegiate football players showed an inverse association between total years of play and left hippocampal volume.[38] The study indicated that the longer athletes had played football the smaller their left hippocampus.[38] A possible link between thalamic volume and total years of play was investigated in publication B.

Age at first exposure

Compared with the total impact burden mentioned above, the age at first exposure (AFE) is reflective of another aspect of exposure to RHI. In terms of contact sports, it is the age when an athlete started playing. For a long time, the general opinion was that children tend to recover quicker from brain trauma than adults due to greater neuroplasticity of the immature brain.[55] More recent data, however, casted doubts on this theory. In fact, more recent studies suggest that children seem to be more prone to a complicated course after brain trauma compared with adults.[56-61] In fact, younger age has been described as a risk factor for prolonged post-concussive recovery.[61,62] Further, younger age was associated with worse outcome and persistent structural alterations in

the brain following head trauma.[38,60,63] Despite this evidence, only four previous studies investigated AFE to RHI. Two of these studies investigated the association between neuropsychiatric symptoms and AFE in a cohort of living athletes. They showed that younger AFE was associated with more later-life impairment in executive function and memory as well as more apathy, depression, and dysregulation of behavior.[50,64] The only previous imaging study on AFE in living athletes showed that younger AFE was associated with alterations in the corpus callosum.[65] More importantly, a recent study by Alosco and colleagues among neuropathologically confirmed cases of CTE showed that younger AFE to contact sports was associated with earlier symptom onset in CTE.[66] The higher vulnerability to head trauma of the immature brain may be closely linked to periods of brain development during childhood and adolescence.[67,68] A more detailed discussion of possible background theories for greater susceptibility of the developing brain is given in publication B.

Moreover, in previous studies further risk factors have been proposed and investigated including position played[69], beta-amyloid pathology[70], management of previous concussions, concomitant psychiatric or neurologic diseases, substance abuse, obesity, age, cognitive reserve[26,48,71,72], as well as biological parameters in particular APOE-ε4 genotype[14,73-75].

1.1.8 Treatment and prevention

A causal treatment for CTE is currently not available. There have been tests in rodent models on inhibition of glycogen-synthase-kinase-3.[76] This enzyme has shown to be a promoter of apoptosis and an inhibitor of neuroprotective cascades and is blocked by other enzymes after traumatic brain injury. Stimulating this blockage may lead to enforcement of the neuroprotective cascade and therefore protect against sequelae of head trauma.[76] Further, an antibody against cis-p-tau which was suggested to be a key initiator of the neuropathological cascade after RHI was introduced by Kondo and colleagues.[25] Although these are important first steps, the search for treatment options of CTE is still in its infancy. In the long run, more research is needed to gain a complete picture and better understanding of etiology, pathogenesis, and modulating factors of CTE.

2. METHODOLOGICAL CONSIDERATIONS

2.1 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) enables to image soft tissues of the body. The underlying physical principle of MRI is nuclear magnetic resonance (NMR).[77] ^{1}H atoms have a spin and therefore NMR property. Spin axes of the atoms usually point randomly into different directions making the sum of all magnetic moments in a spin system, or net magnetization, very small. During an MRI session, a strong magnetic field (B_0) is created in the scanner and aligns the protons' spin axes to the external magnetic field direction. Due to their angular momentum the nuclei rotate or process around the B_0 axis. The velocity of rotation around B_0 is called Larmor frequency.[77]

The nuclei can be excited by application of another radiofrequency magnetic field B_1 perpendicular to B_0. During application of B_1, the nuclei absorb energy. With the end of the B_1 pulse nuclei relax again. The longitudinal relaxation is called the T1-relaxation. The time constant T1 describes the duration of the longitudinal relaxation until 63% of the protons have returned to equilibrium state following an excitement.[78] Depending on the type of tissue, returning times to B_0 equilibrium state vary between protons. For example, protons in fat quickly return to B_0 while protons in water or cerebro-spinal fluid (CSF) slowly return, making fat appear bright and water or CSF dark on a T1-weighted image.[78] At the same time of the longitudinal relaxation, protons de-phase from their transverse precession which is called transverse relaxation or T2-relaxation. The time constant that describes the transverse relaxation is referred to as T2.[78] In general, T1-weighted images generate a good contrast between gray and white matter which is why T1-weighted images are widely used to investigate the morphology and anatomy of the brain.[79] The protons' energy release can be detected by scanner coils and displayed as the "free-induction decay (FID)".[78] The images can be generated on the basis of the FID and through a mathematical function known as Fourier transform.[78]

2.2 Quantification of Brain Regional Volumes

The calculation of brain regional volumes on the basis of a subject's MRI scan plays an essential role in studying the effects of RHI. Numerous previous studies investigating sequelae of RHI have reported brain regional volume reduction (e.g.,[38,39,80]). The importance of volumetric studies is further emphasized by focal brain atrophy found in confirmed cases of CTE.[3] Theoretically, to calculate the volume of a defined region of interest (ROI) in a subject's MRI scan, it is necessary to select all to the ROI corresponding voxels in each slice of the MRI scan. Various information is taken into consideration when selecting the voxels and defining the ROI manually, including overall position in the brain, anatomic landmarks, position relative to adjacent structures, and image intensities.[81] Having selected all corresponding voxels in each slice of the scan, the surface area can be calculated and multiplied by the slice thickness in order to obtain the brain regional volume. Until today, manual tracing is considered to be the gold standard for extracting brain volumes.[82] A major disadvantage of this manual technique, however, is its feasibility. Imaging studies to date mostly involve large number of subjects. Labelling some or all anatomical structures of the brain in one MRI scan may take up to a week for a

trained anatomist.[81] Therefore, the use of automated tools for whole brain segmentation is common in neuroimaging studies. A tool routinely used to investigate brain regional volumes and also used in the studies introduced here is the open-access software FreeSurfer (http://surfer.nmr.mgh.harvard.edu; Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA).

2.3 FreeSurfer

2.3.1 Overview

FreeSurfer is a software designed to analyze MRI data and is well known for segmenting the brain into standardized ROIs. Its essential tools include surface- and volume-based analyses, investigation of cortical thickness, reconstruction of the white matter surface, as well as subsegmentation of the hippocampus.[83] The volume-based analysis stream enables to calculate subcortical and cortical volumes and was used in both original articles introduced in this dissertation. A detailed description of the volume-based analysis stream including its mathematical and technical background is provided in [81,84]. A brief description of the steps comprising the pipeline is provided below.

2.3.2 Pre-Processing

The structural T1-weighted images are skull-stripped with the aid of a brain mask. A motion correction and transformation to Talairach space is performed.[85-87] Further, an initial volumetric labelling is performed and the images are corrected for intensity variations.[81,88,89]

2.3.3 Label Construction and Assignment

For the final parcellation various information is taken into account, that is, intensities and spatial information including local relationships between neighboring voxels and structures.[81,84] For the following steps an atlas is used containing probabilistic information, e.g. on intensities of anatomical classes and locations.[81] The probabilistic atlas is created on the basis of a training set which was labelled manually. One of the key features of the probabilistic atlas is that coordinates have anatomical meaning as opposed to coordinates in raw images which vary by subject-dependent parameters such as position of the head in the scanner.[81] For each voxel the intensity is measured locally and a label is allocated through the probabilistic atlas.[81] The use of intensity information alone is not possible given the great overlap between different structures and tissue type (e.g., gray and white matter overlap by more than 12%).[81] Next, for each voxel specific probabilities are calculated. First, the likelihood that a voxel appears at a certain location in the probabilistic atlas is calculated. Second, the local relationships between different anatomic structures are accounted for in terms of the probability that a certain label is assigned to a certain voxel considering the labels assigned to its neighboring voxels.[81] This way, it is possible to integrate local anatomic information into the parcellation such as the "posterior amygdala is frequently superior to anterior hippocampus, but never inferior to it"[81].

For the whole brain segmentation process all information described above is incorporated. An interim segmentation is computed by matching each voxel to the label for which the likelihood is greatest. The

neighborhood function is used to re-evaluate the probabilities for each voxel and label which leads to a re-segmentation based on the new probabilities. The re-evaluation/re-segmentation process stops and the final parcellation is obtained once the probabilities for each voxel and assigned label stay the same.[81,84]

Fig. 2 Results of FreeSurfer processing stream.

2.3.4 Possible limitations

FreeSurfer is routinely used in neuroimaging studies. It has enabled researchers to segment even large datasets in reasonable amounts of time and has yielded insight into various forms of neuropathology, including sequelae of RHI (e.g., [90-95]). Whenever using automated segmentation tools to subsegment datasets, the results should be validated in terms of anatomical accuracy. However, common standard in research studies is to not perform visual quality assessments but to rely on the automated segmentation results.[96] Whether or not FreeSurfer can be applied unrestrictedly and accurately without manual editing to all forms of neuropathology still remains to be elucidated. The training set used by FreeSurfer to create its subcortical atlas consists of 30 healthy brains and 11 brains with diagnosed dementia [97] which may possibly affect the segmentation results. Given its common use in imaging studies to date, only a surprisingly small number of studies have compared FreeSurfer with the gold standard manual tracing.[82,98-133] Even less studies have compared the segmentation technique in context with clinical symptoms.[104,107] Importantly, several of these studies have shown aberrant results when comparing automated with manual segmentation.[98,99,101-104,106,107,109,111-113,116,117,119-124,126-128,130,131,133] Most of these studies investigated segmentation results of the hippocampus.[82,98,100-113,116-118,120,121,123,126-129,131,133] To a lesser extent, also other anatomical structures have been analyzed, including the temporal lobe and amygdala, basal ganglia, thalamus, parts of the white matter, the insular and frontal cortex, as well as the lateral ventricles.[82,98,99,107,110,114,115,118-120,122,124,125,128,130,132] Significant discrepancies were reported for several structures and in particular for the hippocampus.[98,99,101-104,106,107,109,111-113,116,117,119-124,126-128,130,131,133] Especially in subjects with advanced forms of atrophy or anatomical anomalies, FreeSurfer has shown significantly differing results compared with the gold standard manual tracing.[103,117,121,123] This is of particular note given the frequent use of FreeSurfer in samples with expected focal or global brain atrophy, such as in samples at risk of CTE.

Additionally, it remains unclear how accurate FreeSurfer performs in segmenting anatomical structures other than the named above.

2.4 Neurobehavioral assessment

In the following two studies a neurobehavioral testing battery was administered to the subjects. The raw test results were converted to standard scores controlled for sex, age, and education.[134] Next, test results were grouped to four factor scores based on principal component analysis: mood and behavior, attention/psychomotor speed, visual and verbal memory.[134] The tests included are briefly described below.

Table 1. Neuropsychological testing battery

Mood and Behavior	
Apathy Evaluation Scale	The Apathy Evaluation Scale is an 18-item scale used to quantify and characterize different levels of apathy in adults. It is not syndrome-specific and can be used in various disorders.[135]
Beck Depression Inventory II	The Beck Depression Inventory II is a 21-item instrument to assess the severity of depression in adults. The subject is presented 21 symptoms of depression and has to choose between four different options to answer best how he felt in the preceding two weeks.[136]
Beck Hopelessness Scale	The Beck Hopelessness Scale is a 20-item true/false scale to measure hopelessness in adults.[137]
Barratt Impulsivity Scale	The Barratt Impulsivity Scale is a 30-item questionnaire to assess a subject's impulsiveness. The items display impulsive/non-impulsive behavior rated on a four-point scale.[138]
Behavior Rating Inventory of Executive Functioning – Adult Version	The Behavior Rating Inventory of Executive Functioning – Adult Version is a 75-item measure used to assess an adult's self-report on executive function and self-regulation. It is not

	syndrome-specific and can be used in a variety of disorders.[139]
Center for Epidemiologic Studies – Depression Scale	The Center for Epidemiologic Studies – Depression Scale is a 20-item screening test on depressive symptoms based on the DSM-V definition of a major depressive episode.[140]
Hamilton Depression Rating Scale	The Hamilton Depression Rating Scale is a 21-item rating scale designed to assess the severity of depression in an interview with only the first 17 items being scored.[141]
Brown-Goodwin Lifetime History of Aggression	The Brown-Goodwin Lifetime History of Aggression is an 11-item scale with each item measuring different aspects of aggressive behavior. In the original version each item is rated on a 4-point scale.[142]
Modified Scale for Suicidal Ideation	The Modified Scale for Suicidal Ideation is an 18-item scale designed to assess the presence or absence of suicidal intention and categorizes suicidal ideation and desire from low to severe.[143]
Buss Durkee Inventory	The Buss Durkee Inventory is a 75-item scale used to assess hostility and aggression with a true/false answer design.[144]
Attention/Psychomotor Speed	
Controlled Oral Word Association Test	The Controlled Oral Word Association Test is designed to assess verbal fluency. The number of words starting with “C”, “L”, or “F” an adult can produce in one minute of time are counted. At the end of minute one, the procedure is repeated with

	the next letter, thus, taking three minutes in total.[145]
Delis-Kaplan Executive Function System Color Word Interference Test	The Delis-Kaplan Executive Function System Color-Word Interference Test is a measure to assess executive functioning in children through adults 8-89 years. It consists of four parts: word reading, color naming, inhibition, and inhibition/switching.[146]
Trail Making Test	The Trail Making Test is a measure designed to assess visual attention. It consists of two parts: The subject has to connect numbers in ascending order distributed on a piece of paper, in the second part numbers and letters have to be connected in ascending order. Results are displayed as the number of seconds needed to fulfill the task.[147]
Wechsler Adult Intelligence Scale – Revised	The Wechsler Adult Intelligence Scale – Revised is designed to measure intelligence in adolescents, starting at age 16, and adults. It consists of 11 subtests in total, thereby 6 verbal subtests and 5 non-verbal subtests. A verbal-, performance-, and full IQ score can be obtained.[148]
Verbal Memory	
Neuropsychological Assessment Battery – Story Learning	The Neuropsychological Assessment Battery – Story Learning was designed to assess verbal memory. The subject is told a five-sentence story. There are early as well as delayed recall trials. Content as well as wording are accounted for in the scoring.[149]

Neuropsychological Assessment Battery – List Learning	The Neuropsychological Assessment Battery – List Learning is used to assess verbal episodic memory. The subject is presented 12 words groupable into 3 different semantic categories. First, the subject has to recall as many words as possible in three trials. Then different, partially overlapping words are presented that the subject has to recall. Afterwards, the subject again has to recall the words which were presented at first.[149]

Visual Memory	
Boston Qualitative Scoring System (BQSS) for the Rey-Osterrieth Complex Figure	The Rey-Osterrieth Complex Figure is used to assess visuospatial skills and memory. A complex figure consisting of various elements is presented to the subject. First, the subject has to copy the figure. Next, there is a short recall trial followed by a long recall trial after 30 minutes. The BQSS thereby provides a qualitative assessment of executive skills while producing the figure. [150]

3. RESEARCH QUESTIONS AND HYPOTHESES

The goal of this dissertation is to evaluate gray matter alterations in a sample of symptomatic former professional football players who were exposed to RHI and are at risk of CTE.

In order to address this goal, we chose to conduct volumetric analyses of regional gray matter volumes given the evidence of brain regional volume loss in living athletes as well as neuropathologically confirmed cases of CTE.[3,20,35] As mentioned previously, there are different possibilities to calculate brain regional volumes. Still, the gold standard in research studies is to derive regional volumes manually, however, the number of studies that actually perform manual tracing is sparse. This is due to the time effort and expertise needed for this method. In fact, FreeSurfer has shown promising results in volumetric studies of RHI as well as other forms of neuropathology.[83,90-93] However, to date, only a surprisingly small number of studies have investigated FreeSurfer's reliability and anatomical precision in comparison to manual tracing.[82,98-133] In fact, numerous of these studies reported volumetric or anatomical inaccuracies of FreeSurfer.[98,99,101-104,106,107,109,111-113,116,117,119-124,126-128,130,131,133] To what extent this may influence associations with neurobehavioral features is even less known.[104,107] More importantly, inaccuracies of FreeSurfer were more pronounced when subjects with high volumetric divergence from the norm or anatomical anomalies were analyzed.[103,117,121,123] To our knowledge, no previous study to date has evaluated FreeSurfer's reliability in comparison to manual editing of the results in a sample exposed to RHI and at risk of CTE. Further, while the majority of previous studies have investigated the hippocampus, the precision and reliability of FreeSurfer in labelling other anatomical structures of interest in studies of RHI, such as amygdala, cingulate gyrus, corpus callosum, or the lateral ventricles remains even less clear.

Finding a diagnostic marker for CTE in the living is one of the main aims of current CTE research. Therefore, precise volumetric analyses among living athletes are particularly important given the neuropathologic evidence of subcortical atrophy described in CTE. In a previous study which is not part of this thesis I analyzed volumetric differences of amygdala, hippocampus, and cingulate gyrus among living athletes compared with controls without prior brain trauma. Compared with controls we found reduced volumes among the players group and we further showed that reduced volumes of hippocampus and cingulate gyrus were associated with worse neurobehavioral function. As highlighted previously, the thalamus is a central subcortical structure involved in multiple neuronal networks and functions. Despite its important anatomical function and evidence from confirmed *postmortem* cases of CTE, to date, only Bernick and colleagues investigated the association between exposure to RHI and thalamic volume *in vivo*.[36,37,43] They included a sample of professional boxers and mixed martial arts fighters found a reduction of thalamic volumes with every fight and with total years of fighting.[36,43] Although head accelerations in professional football are commonly lower than in professional fighting[44,45] we hypothesized to see a similar effect in a sample of former professional American Football players.

To date, exposure to RHI is the most important risk factor for CTE.[1-3] However, to gain a better picture of CTE and in the long run protect those at risk, identifying further possible risk factors is crucial. There has been a shift in the field towards the theory that the immature brain has a higher vulnerability towards brain trauma.[38,60,61] In line with this, three previous *in vivo* studies as well as one *postmortem* study suggested AFE to contact sports

as a possible risk factor for CTE.[50,64-66] Moreover, younger AFE to contact sports has been linked to earlier symptom onset of CTE in neuropathologically confirmed cases.[66]

Therefore, in the following presented studies, we intended to

1) Evaluate the accuracy of FreeSurfer and compare the results with manually edited volumes

2) Apply FreeSurfer to derive thalamic volumes and explore the association between measures of exposure to RHI, in particular AFE to American Football,

in a sample of symptomatic former professional football players who were exposed to RHI routinely during their career and are at risk of CTE.

4. PUBLICATION A

4.1 Background

The original article A with the title "Automated Versus Manual Segmentation of Brain Region Volumes in Former Football Players" was published in April 2018 in *Neuroimage: Clinical* (IF: 3.94; citations: 9). In this study we investigated whether or not the automated segmentation of 11 brain regions often shown to be affected in studies of RHI can be used reliably without manual correction in a sample of symptomatic former professional football players and age-matched controls without a history of brain trauma. In order to further integrate the results in a clinical context, we also explored the associations between volumes and neurobehavioral measures among the players' group.

4.2 Methods

Eighty-six former professional football players (mean age = 55.2 ± 8.0 years) and 22 control subjects (mean age = 57.0 ± 6.6 years) were included in this study. The inclusion criteria for former professional football players were: male, 40-69 years of age, at least 12 years of organized football experience with a minimum of 2 years of play in the NFL, self-reported deficits in cognition, mood, and behavior at least 6 months prior to study start. The inclusion criteria for the control subjects were: male and no reported history of participation in organized contact-sports or traumatic brain injury. For all subjects T1-weighted images were acquired on a 3T scanner (Verio, Siemens Healthcare, Erlangen, Germany) and automated FreeSurfer 5.3 segmentation and manual correction of the following 11 brain regions were obtained: bilateral corpus callosum, cingulate gyrus, amygdala, hippocampus, amygdala-hippocampal complex, and lateral ventricles. Manual correction of the FreeSurfer labels was done by three trained individuals under the supervision of a neuroanatomist and a radiologist, respectively. Further, two individuals randomly chose 10 label maps to edit to evaluate inter-observer reliability. Automated segmentation and manually-corrected volumes were compared using an intraclass correlation coefficient (ICC). Mood and behavior, attention/psychomotor speed, and visual and verbal memory were tested with a neuropsychological testing battery. All statistical analyses were done with Statistical Analysis System (SAS version 9.4; SAS Institute Inc., North Carolina, USA). Mixed effects regression models were computed for between-group mean volume comparisons and to correlate former football player regional brain volumes with neurobehavioral test results. Written informed consent was obtained from all participants prior to enrollment.

4.3 Results

The inter-observer reliability was good to excellent among all brain regions (ICC range: 0.723 to 0.990). Further, in this study the ICC was high between automated and manually corrected volumes of the corpus callosum (0.911), lateral ventricles (right 0.980, left 0.967), and amygdala-hippocampal complex (right 0.713, left 0.731). Automated segmentation volumes of the cingulate gyrus (right 0.639, left 0.351) as well as amygdala (right -0.170, left -0.090) and hippocampus (right 0.539, left 0.637) when delineated separately demonstrated low correlation with manually corrected volumes. In general, the mean volume differences between the two groups were greater when comparing the manually corrected volumes than the automatically segmented volumes.

Statistically significant differences between the former football players and controls were found in 8 of 11 manually corrected regions versus only 4 of 11 segmented regions without manual edit (see **Table 2**). There was no concordance among the correlations between the manually edited and segmented brain regions without manual edit and neurobehavioral test results. Among the former professional football players there were 3 statistically significant correlations between brain regional volumes and neurobehavioral factors with the automatically segmented volumes, however, 3 different brain regions and neurobehavioral factors correlated if the manually edited volumes were used (manual correction: cingulate gyrus and attention/psychomotor speed (left: $p = 0.003$; right: $p = 0.005$); left ventricle and visual memory ($p = 0.047$); automated: left amygdala ($p = 0.036$), left hippocampus ($p = 0.036$), and left amygdala-hippocampus-complex ($p = 0.014$) and visual memory).

Table 2. Mean differences of volume between NFL players and controls in mL^3 *(p)* adapted from Publication A, 151

Region	Manual Correction	Automated
Corpus callosum	121.96 (0.265)	129.98 (0.285)
Left amygdala	176.22 (0.005)	159.07 (0.001)
Right amygdala	157.43 (0.012)	165.63 (0.018)
Left hippocampus	158.44 (0.024)	207.12 (0.087)
Right hippocampus	146.09 (0.032)	86.03 (0.535)
Left amygdala-hippocampal complex	315.43 (0.031)	366.19 (0.007)
Right amygdala-hippocampal complex	300.88 (0.031)	251.67 (0.045)
Left cingulate gyrus	575.01 (0.036)	286.12 (0.535)
Right cingulate gyrus	475.28 (0.032)	80.14 (0.827)
Left ventricle	2300.38 (0.155)	1138.77 (0.474)
Right ventricle	2619.67 (0.118)	1382.15 (0.376)

4.4 Discussion

The automated segmentation performed with FreeSurfer 5.3 shows excellent correlations between manually edited volumes of the corpus callosum and the lateral ventricles and the segmented volumes without manual correction. However, results from correlations between manually edited volumes of amygdala, hippocampus, and cingulate gyrus and segmented volumes without editing appeared to be weaker. Moreover, the associations between brain regional volumes and neurobehavioral factors in this sample of former professional football players have pointed entirely different results when using the automatically segmented volumes and the manually corrected volumes. When computing group comparisons, more differences were identified when using the manually corrected volumes than the automatically derived ones possibly due to greater regard to inter-individual differences when no atlas was used. Findings from this study suggest that the cingulate gyrus should undergo manual correction prior to data analysis, however, the literature on manual segmentation results for the cingulate gyrus is very sparse.[152] Therefore, future studies are needed to confirm this result. The finding that FreeSurfer performs less accurate on segmenting the hippocampus and amygdala compared with manual tracing is in line with the literature.[98,99,101-104,106,107,109,111-113,116-118,120,121,123,126-128,131,133] In particular, the majority of studies that have investigated the two structures have shown an overestimation of volumes compared with

manual tracing.[99,104,109,111-113,116,120,126-128,131] A possible reason for this segmentation inaccuracy of these medial temporal lobe structures may be partial volume effects that lead to inclusion of more voxels in automated segementation.[128] Another reason may be that the contrast between the two structures on a T1-weighted image is not optimized for this question. Delineation is further aggravated by the fact that there may be inter-individual shape or volumetric differences depending on several circumstances including the neuropathology or age-group studied. In fact, evidence exists that T2-weighted imaging of the hippocampus outperforms T1-weighted imaging especially in rating stages of hippocampal atrophy.[153,154] Better matches of results were achieved in this study when amygdala and hippocampus volumes were included as an amygdala-hippocampal complex. The combination may mitigate the previously mentioned issues; however, it should be taken into consideration that amygdala and hippocampus are associated with two independent memory systems.[155] While the hippocampus has a major functional role in episodic memory, the amygdala is majorly specialized in processing emotional memories and conditioning of fear.[155] Still, episodic memory storage may be influenced by emotional perceptions and, thus, amygdala and hippocampal-dependent memories may be influenced by each other.[155] Therefore, findings from this study indicate that amygdala and hippocampus may be analyzed as amygdala-hippocampal complex or should undergo manual correction prior to data analysis depending on the research question.

The results of this study may not be generalizable to other versions of FreeSurfer. Also, the findings may not be transferrable to studies of shape or volume overlap. The results were obtained in a sample of symptomatic former professional football players who were exposed to RHI; therefore, the findings may not transfer to studies of different pathologies.

4.5 Contribution

This study is part of my work I conducted during one year at the Psychiatry Neuroimaging Laboratory in Boston, USA, at this worldwide unique dataset. My contribution to this study consisted of the following:

1. Data analysis and interpretation together with the first author.
2. Critical review of the manuscript.

NeuroImage: Clinical 18 (2018) 888–896

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Automated versus manual segmentation of brain region volumes in former football players

Jeffrey P. Guenette[a,b], Robert A. Stern[c,d], Yorghos Tripodis[c,e], Alicia S. Chua[e], Vivian Schultz[a,f], Valerie J. Sydnor[a], Nathaniel Somes[a], Sarina Karmacharya[a], Christian Lepage[a], Pawel Wrobel[a,f], Michael L. Alosco[c], Brett M. Martin[g], Christine E. Chaisson[c,g], Michael J. Coleman[a], Alexander P. Lin[a,b,h], Ofer Pasternak[a,b], Nikos Makris[a,b,i], Martha E. Shenton[a,b,j], Inga K. Koerte[a,f,*]

[a] Psychiatry Neuroimaging Laboratory, Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States
[b] Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States
[c] BU Alzheimer's Disease and CTE Center, Boston University, Boston, MA, United States
[d] Departments of Neurology, Neurosurgery, and Anatomy & Neurobiology, Boston University School of Medicine, Boston, MA, United States
[e] Department of Biostatistics, Boston University School of Public Health, Boston, MA, United States
[f] Department of Child and Adolescent Psychiatry, Psychosomatic, and Psychotherapy, Ludwig-Maximilian-University, Munich, Germany
[g] Data Coordinating Center, Boston University School of Public Health, Boston, MA, United States
[h] Center for Clinical Spectroscopy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States
[i] Center for Neural Systems Investigations, Massachusetts General Hospital, Boston, MA, United States
[j] VA Boston Healthcare System, Brockton Division, Brockton, MA, United States

ARTICLE INFO

Keywords:
Magnetic resonance imaging
Neuroimaging
Head trauma
Traumatic brain injury
Chronic traumatic encephalopathy

ABSTRACT

Objectives: To determine whether or not automated FreeSurfer segmentation of brain regions considered important in repetitive head trauma can be analyzed accurately without manual correction.

Materials and methods: 3 T MR neuroimaging was performed with automated FreeSurfer segmentation and manual correction of 11 brain regions in former National Football League (NFL) players with neurobehavioral symptoms and in control subjects. Automated segmentation and manually-corrected volumes were compared using an intraclass correlation coefficient (ICC). Linear mixed effects regression models were also used to estimate between-group mean volume comparisons and to correlate former NFL player brain volumes with neurobehavioral factors.

Results: Eighty-six former NFL players (55.2 ± 8.0 years) and 22 control subjects (57.0 ± 6.6 years) were evaluated. ICC was highly correlated between automated and manually-corrected corpus callosum volumes (0.911), lateral ventricular volumes (right 0.980, left 0.967), and amygdala-hippocampal complex volumes (right 0.713, left 0.731), but less correlated when amygdalae (right −0.170, left −0.090) and hippocampi (right 0.539, left 0.637) volumes were separately delineated and also less correlated for cingulate gyri volumes (right 0.639, left 0.351). Statistically significant differences between former NFL player and controls were identified in 8 of 11 regions with manual correction but in only 4 of 11 regions without such correction. Within NFL players, manually corrected brain volumes were significantly associated with 3 neurobehavioral factors, but a different set of 3 brain regions and neurobehavioral factor correlations was observed for brain region volumes segmented without manual correction.

Conclusions: Automated FreeSurfer segmentation of the corpus callosum, lateral ventricles, and amygdala-hippocampus complex may be appropriate for analysis without manual correction. However, FreeSurfer segmentation of the amygdala, hippocampus, and cingulate gyrus need further manual correction prior to performing group comparisons and correlations with neurobehavioral measures.

* Corresponding author at: Psychiatry Neuroimaging Laboratory, 1249 Boylston St., Boston, MA 02215, United States.
E-mail address: ikoerte@bwh.harvard.edu (I.K. Koerte).

https://doi.org/10.1016/j.nicl.2018.03.026
Received 9 November 2017; Received in revised form 2 February 2018; Accepted 21 March 2018
Available online 21 March 2018

1. Introduction

Analysis of brain regional volumes has yielded insight into the pathology and pathophysiology of a variety of neurological and psychiatric diseases including Alzheimer's disease (see reviews by Kantarci and Jack, 2003 and Busatto et al., 2008), schizophrenia (see metaanalysis by Olabi et al., 2011 and reviews by Hulshoff Pol and Kahn, 2008 and Shenton et al., 2010), post-traumatic stress disorder (see reviews by Ahmed-Leitao et al., 2016 and Milani et al., 2017), mild traumatic brain injury (see reviews by Shenton et al., 2012 and Mu et al., 2017) and repetitive head trauma (see reviews by Ng et al., 2014 and Koerte et al., 2015), to name just a few. Accurate and precise volumetric measurements are essential for both reliability and reproducibility. Given the time-consuming nature of manual segmentation, automated segmentation techniques are critical for studies involving large imaging datasets. Moreover, to be useful in the clinical setting, automated segmentation techniques are also critical given that time-consuming manual segmentation by a radiologist for interpretation is not feasible. However, in addition to segmenting the brain in a short period of time, automated segmentation must also provide levels of accuracy and precision that yield results similar to those obtained with manual segmentation, which is currently the gold standard.

Although some automated segmentation algorithms have shown potentially promising results (see review by Dill et al., 2015), many often provide suboptimal results (e.g. de Flores et al., 2015; González-Villà et al., 2016; Grimm et al., 2015; Haller et al., 2016; Næss-Schmidt et al., 2016; Schoemaker et al., 2016) and there is thus ongoing research to develop better algorithms (Akhondi-Asl et al., 2011; Inglese et al., 2015; Mendrik et al., 2015).

Neuroimaging volumetry studies routinely utilize freely-available automated segmentation tools such as FreeSurfer (http://surfer.nmr.mgh.harvard.edu; Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA). Some studies of FreeSurfer have shown deficiencies in automated segmentation of the cerebral cortex (Makris et al., 2008), hippocampus (Cherbuin et al., 2009; de Flores et al., 2015; Grimm et al., 2015; Morey et al., 2009; Wenger et al., 2014), and amygdala (Grimm et al., 2015; Morey et al., 2009; Schoemaker et al., 2016), but data regarding the accuracy and precision of FreeSurfer is not readily available for other important and frequently studied regions including the cingulate gyrus, corpus callosum, and lateral ventricles, all areas important in the investigation of repetitive head trauma. Moreover, there are no published data that demonstrate whether study outcome measures are concordant or discordant when using automated segmentation as compared to manual segmentation.

Volumetric analysis of the brain is particularly important in individuals with exposure to repetitive head trauma as there is evidence that repetitive head impacts may result in regional brain atrophy (Bernick et al., 2015; Goddeyne et al., 2015; Laurent et al., 2010; McKee et al., 2009). Players of American football have a particularly high exposure to repetitive head impacts. For example, college American football players sustain a median of 420 head impacts per season and some players sustain over 2400 head impacts per season, as measured by accelerometers (Crisco et al., 2011).

The aim of this study was to determine whether or not FreeSurfer automated segmentation can be used reliably, without the need for manual brain volume editing, in studies of repetitive head impact that investigated the volumes of the cingulate cortex (left and right), corpus callosum, amygdala (left and right), hippocampus (left and right), amygdala-hippocampal complex (left and right), and lateral ventricles (left and right) in retired National Football Players (NFL) and same aged controls without history of contact sports or brain injury.

2. Methods

This study utilized data from the Diagnosing and Evaluating Traumatic Encephalopathy using Clinical Tests (DETECT) study, funded by the National Institutes of Health (NIH). The DETECT study details have been described in prior publications (Alosco et al., 2016, 2017; Stamm et al., 2015; Stern et al., 2016). All study procedures were approved by the Boston University Medical Center Institutional Review Board and all neuroimaging procedures were approved by the Partners Institutional Review Board. All subjects provided written, informed consent.

2.1. Participants and procedure

There were two cohorts in the DETECT study: former NFL players with at least 12 years of organized football experience, at least 2 years of active participation in the NFL, and self-reported declines in cognition, mood, and behavior within 6 months of study commencement; and control subjects with no reported history of participation in organized contact sports or traumatic brain injury. All subjects were male, aged 40 to 69 years, spoke English as their first language, had no contraindication to MR imaging or lumbar puncture, and no history or diagnosis of central nervous system (CNS) disease.

Of the 96 enrolled former NFL player subjects, 10 were excluded due to inadequate or absent neuroimaging data, resulting in a final sample size of 86 former NFL players (age: 55.2 ± 8.0 years). Of these 86 subjects, complete neurobehavioral testing results were available for a total of 76 subjects. Neuroimaging data was available for all 28 control group subjects, 3 of whom were excluded due to image quality and 3 more were excluded due to subsequently identified CNS disease, contact sport participation, or history of mild traumatic brain injury, resulting in a final sample size of 22 control subjects (age: 57.0 ± 6.6 years).

All subjects were evaluated according to the DETECT neurobehavioral and neuroimaging protocol, including neuroimaging, structured psychiatric interview, and neuropsychological testing.

2.2. MRI data acquisition

DETECT neuroimaging was performed at Brigham and Women's Hospital on a 3-Tesla MRI system (Verio, Siemens Healthcare, Erlangen, Germany) with a 32-channel head array and the Syngo MR-B17 software suite. Only the T1-weighted magnetization prepared rapid gradient echo (TR = 1800 ms, TI = 1100 ms, TE = 3.36 ms, voxel size = 1 × 1 × 1 mm, acquisition matrix = 256 × 256, flip angle = 7°) sequence was used for this study.

2.3. Image processing

All T1-weighted images were visually inspected for quality. Brain masks of each subject were generated by FreeSurfer 5.3 (http://surfer.nmr.mgh.harvard.edu; Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA) and corrected manually. Each brain was segmented using T1-weighted images and FreeSurfer 5.3. This process yielded label maps of deep gray matter, white matter, and CSF structures (including the hippocampus, amygdala, corpus callosum, and lateral ventricles). This process also yielded parcellation label maps of the cerebral cortex (including the cingulate gyrus) based on gyral and sulcal structures. The FreeSurfer option for utilizing T2 or FLAIR image contrast to improve pial surface estimations along CSF borders was not used for this study. Estimated total intracranial volumes were also calculated using the automated FreeSurfer method (Buckner et al., 2004).

FreeSurfer segmentation and parcellation maps were then loaded into the Editor module of Slicer 4.5.0 (http://www.slicer.org, Surgical Planning Laboratory, Brigham and Women's Hospital, Boston, Massachusetts, USA) (Fedorov et al., 2012) and overlayed on the aligned T1-weighted images with image interpolation turned off.

Following written directions based on the below-described

approaches for each region, two trained raters manually corrected the image label maps of the amygdala, hippocampus, corpus callosum, and cingulate gyrus. One rater corrected approximately two-thirds of the cases while the second rater corrected the remainder. All cases were then reviewed by a single neuroanatomist for accuracy. A third trained rater manually corrected the image label maps of the lateral ventricles and a radiologist reviewed the corrected lateral ventricle label maps for accuracy. All investigators and raters were blind to group membership at the time of segmentation and review by a trained expert.

To evaluate inter-observer reliability, a fourth trained rater, following the same written directions, corrected the FreeSurfer image label maps of the amygdala, hippocampus, corpus callosum, and cingulate gyrus in 10 randomly chosen subjects, while a radiologist corrected the FreeSurfer image label maps of the lateral ventricles in 10 randomly chosen subjects. A neuroanatomist reviewed the label maps corrected by the trained raters. These individuals were blind to the previously corrected label maps and to group membership at the time of segmentation.

2.3.1. Amygdala and hippocampus

Amygdala and hippocampus volumes were manually corrected based on an approach described by Gurvits et al., 1996. Label maps were corrected on coronal slices, from anterior to posterior, using sagittal slices for verification. Close attention was given to the anterior portion of the amygdala at the height of the frontotemporal junction as these parts were variably included by FreeSurfer. The posterior boundary of the amygdala was defined as the last coronal slice before the appearance of the mammillary bodies. The anterior border of the hippocampus was defined as the image slice where the mammillary bodies first appeared. The posterior boundary of the hippocampus was defined as the coronal slice where the crus of the fornix was last seen. The amygdala and hippocampus were defined and strictly separate volume entities and therefore volume overlap between the structures was not allowed.

Although the hippocampus and amygdala are typically evaluated independently, the posterior aspect of the amygdala partially overlaps anatomically with the hippocampal head (Kiernan, 2012) and the boundary between the structures is generally not visible below the level of the temporal horn (Chera et al., 2009). As such, the amygdala-hippocampal complex was evaluated as an additional variable by adding the volumes of the two regions together for the purpose of analyzing hypothesized FreeSurfer imprecision in discriminating the boundary of these regions.

2.3.2. Cingulate gyrus

The cingulate gyrus was defined as the gyrus superior to the corpus callosum, identified primarily on sagittal images (Wible et al., 1997). The medial aspect of the cingulate gyrus was confirmed on sagittal images by identifying the callosomarginal fissure on a paramidsagittal slice. Then, working from medial to lateral on sagittal images, and using coronal images for verification, voxels extending to the corpus callosum, paracingulate gyrus, or beyond the rostrum of the corpus callosum into Brodmann's Area 25 were excluded.

2.3.3. Corpus callosum

Given high contrast between the high T1 signal of the corpus callosum relative to surrounding tissues, most borders of the corpus callosum were obvious. As drawn by FreeSurfer, the corpus callosum label maps were confined to the 5 most midline sagittal slices.

2.3.4. Lateral ventricles

Given high contrast between the low T1 signal of the ventricles and surrounding brain, most borders of the lateral ventricles were obvious. Working superiorly to inferiorly on the axial images, and using the coronal and sagittal images for confirmation, separate volumes were drawn for the right and left lateral ventricles, respectively. When the septum pellucidum was not clearly visible, a midline strip of 1 voxel in width was not assigned to either ventricle to ensure separation of the volumes. The volumes were terminated at the foramina of Monro and choroid plexus was excluded. Any voxels identified by FreeSurfer as lateral ventricle but extending into the foramina of Monro, third ventricle, or ambient cisterns were removed.

2.4. Neurobehavioral measures

As part of the DETECT protocol, data from the following neurobehavioral tests and self-report measures were converted to age-, gender-, and education-standardized scores and grouped into four factor scores, based on principal component analyses (Alosco et al., 2016). *Mood and Behavior* included the following tests: Apathy Evaluation Scale, Beck Depression Inventory II, Beck Hopelessness Scale, Barratt Impulsivity Scale, Behavior Rating Inventory of Executive Functioning - Adult Version, Center for Epidemiologic Studies - Depression Scale, Hamilton Depression Rating Scale, and the Brown-Goodwin Lifetime History of Aggression. *Attention and Psychomotor Speed* included: Controlled Oral Word Association Test, Delis-Kaplan Executive Function System Color Word Interference Test (inhibition/switching), Trail Making Test, and the Wechsler Adult Intelligence Scale – Revised Digit Symbol test. *Verbal Memory* included: Neuropsychological Assessment Battery (NAB) Story Learning (phrase unit immediate and delayed recall), and NAB List Learning (short and long delayed recall). *Visual Memory* included: Boston Qualitative Scoring System for the Rey-Osterrieth Complex Figure (immediate recall presence and accuracy and delayed recall presence and accuracy).

2.5. Statistical analyses

Statistical Analysis System (SAS) software (SAS version 9.4; SAS Institute Inc., North Carolina, USA) was used for all statistical analyses. Significance was set at a p-value below 0.05.

2.5.1. Part 1: inter-observer reliability of manually-corrected volumes

Intraclass correlation coefficients (ICC) were calculated to compare the two sets of manually-edited segmentation label map volumes for each region. Levels of correlation as classically described by Cicchetti (1994): low with a correlation coefficient below 0.40, fair from 0.40 to 0.59, good from 0.60 to 0.74, and excellent from 0.75 to 1.00.

2.5.2. Part 2: intraclass correlation of automated and manually-corrected segmentation volumes

Intraclass correlations (ICC) were calculated to compare the automated and manually-corrected segmentation label map volumes for each region. Scatter plots revealed a single outlier subject and that subject was excluded from correlation analysis of the lateral ventricles.

2.5.3. Part 3: brain volume differences between former NFL players and controls using automated segmentation with versus without manual correction

Linear mixed effects regression models were used to estimate between-group volume comparisons (Former NFL Players vs Controls) for the amygdalae, hippocampi, corpora callosa, cingulate gyri, and lateral ventricles. Two regression models were used: 1) a regression model with multivariate outcomes being the amygdalae, hippocampi, and cingulate gyri; and 2) a regression model with multivariate outcomes being the corpus callosum, amygdala-hippocampal complex, and lateral ventricles. All models were controlled for age, body-mass index (BMI), and estimated total intracranial volume (eTIV).

2.5.4. Part 4: association of former NFL player brain volumes with neurobehavioral factors using automated segmentation with versus without manual correction

A linear mixed-effects regression model was performed for all brain

Table 1
Intra-class correlation values of inter-observer reliability for manually-corrected volumes.

Brain region	ICC (95% CI)
Corpus callosum	0.956 (0.864–0.989)
Left amygdala	0.715 (0.264–0.913)
Left hippocampus	0.764 (0.345–0.926)
Left amygdala-hippocampal complex	0.863 (0.576–0.959)
Left cingulate cortex[a]	0.574 (0.011–0.858)
Left ventricle	0.990 (0.965–0.997)
Right amygdala	0.887 (0.656–0.968)
Right hippocampus	0.627 (0.104–0.881)
Right amygdala-hippocampal complex	0.919 (0.741–0.977)
Right cingulate cortex	0.856 (0.576–0.959)
Right ventricle	0.989 (0.930–0.994)

[a]A single outlier in the left cingulate cortex skews the results. Excluding the outlier results in an ICC of 0.724 (0.230–0.919), which is comparable to the ICC for the right cingulate gyrus.
CI = confidence interval.
Dark gray background = excellent correlation.
Light gray background = good correlation.
Levels of correlation as classically described by Cicchetti (1994): low with a correlation coefficient below 0.40, fair from 0.40 to 0.59, good from 0.60 to 0.74, and excellent from 0.75 to 1.00.

regions, excluding the lateral ventricles, to compare segmented brain volumes with grouped neurobehavioral factors. A false discovery rate adjustment was then calculated to account for multiple comparisons. A separate linear regression model was performed for the lateral ventricles due to the typical inverse relationship of lateral ventricle size with brain volumes where ventricle size gets larger as brain parenchyma gets smaller with atrophy (Barron et al., 1976). The neurobehavioral factor models were controlled for age, BMI, eTIV, and years of education.

3. Results

3.1. Part 1: inter-observer reliability of manually-corrected volumes

There was excellent inter-observer correlation for the corpus callosum (ICC = 0.956), left hippocampus (ICC = 0.764), left and right amygdala-hippocampal complex (ICC = 0.863, 0.919), left and right lateral ventricle (ICC = 0.990, 0.989), right amygdala (ICC = 0.887), and right cingulate cortex (ICC = 0.856). Good inter-observer correlation was present in all other regions except for the left cingulate cortex (ICC = 0.574) (Table 1), however there was an outlier identified with left cingulate cortex volumes measuring over 2 standard deviations from the mean and if that outlier was removed then the ICC improved to 0.723. Fair or poor inter-observer correlation was not present in any region.

3.2. Part 2: intraclass correlation of automated and manually-corrected segmentation volumes

The correlation coefficients were high between automated and manually-corrected segmentation volumes for corpus callosum (ICC = 0.911) and right and left lateral ventricles (ICC = 0.977, 0.964). Good correlation was demonstrated with the left and right hippocampus (ICC = 0.637, 0.539), left and right amygdala-hippocampal complex (ICC = 0.731, 0.713) and right cingulate cortex (ICC = 0.639). Lower correlation was demonstrated for the left and right amygdala (ICC = −0.090, −0.170) and for the left cingulate cortex (ICC = 0.351) (Table 2). The correlation of automated and manually-corrected volumes was better for the amygdala-hippocampal complex than for either the amygdala or hippocampus alone, indicating that FreeSurfer better estimates the volume of the amygdala-hippocampal complex than either structure independently.

Table 2
Intra-class correlation values of manually-corrected versus automatically-generated volumes.

Brain region	ICC (95% CI)
Corpus callosum	0.911 (0.871–0.937)
Left amygdala	−0.090 (−0.0272–0.098)
Left hippocampus	0.637 (0.515–0.739)
Left amygdala-hippocampal complex	0.731 (0.629–0.807)
Left cingulate cortex	0.351 (0.175–0.504)
Left ventricle	0.967 (0.956–0.979)
Right amygdala	−0.170 (−0.346–0.017)
Right hippocampus	0.539 (0.393–0.660)
Right amygdala-hippocampal complex	0.713 (0.603–0.792)
Right cingulate cortex	0.639 (0.515–0.739)
Right ventricle	0.980 (0.971–0.986)

Dark gray background = excellent correlation ICC full cohort.
Light gray background = good correlation ICC full cohort.
Levels of correlation as classically described by Cicchetti (1994): low with a correlation coefficient below 0.40, fair from 0.40 to 0.59, good from 0.60 to 0.74, and excellent from 0.75 to 1.00.

3.3. Part 3: brain volume differences between former NFL players and controls using automated segmentation with versus without manual correction

There was generally a larger difference between the mean former NFL player and mean control subject brain region volumes when using manually-corrected volumes than when using uncorrected automated segmentation volumes (Table 3), indicating that there is greater

Table 3
Brain volume comparisons of former NFL players and controls using automated segmentation with versus without manual correction.

Region	Mean difference of volume between former NFL players and controls in mL^3 (p)*	
	With manual correction	Without manual correction
Corpus callosum	121.96 (0.265)	129.98 (0.285)
Left amygdala	176.22 (0.005)	159.07 (0.001)
Left hippocampus	158.44 (0.024)	207.12 (0.087)
Left amygdala-hippocampal complex	315.43 (0.031)	366.19 (0.007)
Left cingulate cortex	575.01 (0.036)	286.12 (0.535)
Left ventricle	2300.38 (0.155)	1138.77 (0.474)
Right amygdala	157.43 (0.012)	165.63 (0.018)
Right hippocampus	146.09 (0.032)	86.03 (0.535)
Right amygdala-hippocampal complex	300.88 (0.031)	251.67 (0.045)
Right cingulate cortex	475.28 (0.032)	80.14 (0.827)
Right ventricle	2619.67 (0.118)	1382.15 (0.376)

*p-Value adjusted for multiple comparisons except for corpus callosum and lateral ventricles as those correlations were calculated in separate regression models.
Dark gray background = significant at $p < 0.05$.

variance in the manually-corrected brain volumes, a finding that is addressed in the discussion. There were also a larger number of statistically significant brain region volume differences between Former NFL Players and Controls when using manually-corrected volumes than when using uncorrected automated segmentation volumes (Table 3), indicating that comparisons performed without manual correction of brain volumes would not have identified all significant group differences in this study.

3.4. Part 4: association of former NFL player brain volumes with neurobehavioral factors using automated segmentation with versus without manual correction

No unedited or manually-corrected volumes were associated with the Mood and Behavior factor score. The manually-edited volumes of the cingulate cortex on the left and right were both associated with the Attention and Psychomotor Speed factor score (left: effect size = 31.053, $p = 0.003$; right: effect size = 25.730, $p = 0.003$), but no unedited volumes were associated with this neurobehavioral factor. No unedited or manually-corrected volumes were associated with the Verbal Memory factor score. The manually-corrected volume of the left ventricle was associated with the Visual Memory factor score (effect size = 3.870, $p = 0.047$), however this statistic would not likely hold under multiple comparisons, while the unedited volumes of the left amygdala, left hippocampus, and left amygdala-hippocampal complex were also associated with the Visual Memory score (amygdala: effect size = 0.798, $p = 0.036$; hippocampus: effect size = 2.617, $p = 0.036$; amygdala-hippocampus complex: effect size = 4.147, $p = 0.014$). There was thus no concordance between which manually-corrected volumes and which unedited volumes reached significance when correlating volumes with neurobehavioral factors (Table 4).

Table 4
Association of former NFL Player brain volumes with neurobehavioral factors using automated segmentation with versus without manual correction.

	Effect size with (-value)* determining significance of NFL player brain volume association with neurobehavioral factor score	
	Volumes with manual correction	Volumes without manual correction
Mood and behavior		
Brain region		
Corpus callosum	−0.335 (0.808)	−0.872 (0.674)
Left amygdala	−1.232 (0.270)	−0.629 (0.173)
Left hippocampus	−0.596 (0.591)	−1.024 (0.515)
Left amygdala-hippocampal complex	−2.350 (0.270)	−2.145 (0.206)
Left cingulate cortex	−5.776 (0.591)	2.377 (0.746)
Left ventricle	1.477 (0.418)	1.536 (0.331)
Right amygdala	−1.253 (0.270)	−1.063 (0.173)
Right hippocampus	−0.870 (0.526)	−1.095 (0.543)
Right amygdala-hippocampal complex	−2.632 (0.270)	−2.543 (0.173)
Right cingulate cortex	−4.199 (0.591)	4.288 (0.674)
Right ventricle	1.865 (0.502)	1.988 (0.473)
Attention and psychomotor speed		
Brain region		
Corpus callosum	2.878 (0.182)	2.976 (0.326)
Left amygdala	0.098 (0.993)	0.032 (0.975)
Left hippocampus	1.903 (0.154)	1.389 (0.472)
Left amygdala-hippocampal complex	2.476 (0.265)	1.584 (0.479)
Left cingulate cortex	31.053 (0.003)	22.880 (0.109)
Left ventricle	1.643 (0.563)	1.364 (0.622)
Right amygdala	0.009 (0.993)	1.194 (0.326)
Right hippocampus	2.109 (0.132)	−0.049 (0.975)
Right amygdala-hippocampal complex	2.628 (0.265)	1.543 (0.479)
Right cingulate cortex	25.730 (0.005)	13.305 (0.463)
Right ventricle	1.224 (0.689)	1.123 (0.672)
Verbal memory		
Brain region		
Corpus callosum	−0.515 (0.951)	−0.601 (0.895)
Left amygdala	−0.266 (0.951)	0.143 (0.895)
Left hippocampus	0.090 (0.951)	0.357 (0.895)
Left amygdala-hippocampal complex	−0.239 (0.951)	0.624 (0.895)
Left cingulate cortex	3.929 (0.951)	0.110 (0.988)
Left ventricle	−0.549 (0.653)	−0.394 (0.635)
Right amygdala	−0.413 (0.951)	0.079 (0.988)
Right hippocampus	0.343 (0.951)	1.187 (0.895)
Right amygdala-hippocampal complex	−0.092 (0.951)	1.336 (0.895)
Right cingulate cortex	−0.934 (0.951)	−4.164 (0.895)
Right ventricle	−0.983 (0.792)	−0.923 (0.846)
Visual memory		
Brain region		
Corpus callosum	0.957 (0.785)	1.275 (0.469)
Left amygdala	0.478 (0.785)	0.798 (0.036)
Left hippocampus	1.274 (0.405)	2.617 (0.036)
Left amygdala-hippocampal complex	2.193 (0.405)	4.147 (0.014)
Left cingulate cortex	3.734 (0.817)	2.550 (0.748)
Left ventricle	3.870 (0.047)	3.776 (0.058)
Right amygdala	−0.123 (0.887)	0.810 (0.279)
Right hippocampus	2.236 (0.083)	1.924 (0.241)
Right amygdala-hippocampal complex	2.636 (0.405)	3.056 (0.069)
Right cingulate cortex	−2.457 (0.821)	−9.957 (0.317)
Right ventricle	4.861 (0.098)	4.123 (0.098)

All p-values adjusted for multiple comparisons except for corpus callosum and lateral ventricles as those correlations were calculated in separate regression models.
Dark gray background = significant at $p < 0.05$.
Note: adjustments for multiple comparisons may lead to slightly different p-values in each study depending on the number of comparisons and the method of adjustment.

4. Discussion

The aim of this study was to determine whether or not FreeSurfer automated segmentation can be used accurately and reliably, without the need for manual brain volume editing, in studies of repetitive head impact that examine the cingulate cortex, corpus callosum, amygdala, hippocampus, and/or lateral ventricles.

4.1. Corpus callosum and lateral ventricles

In this study, automated segmentation volumes of the lateral ventricles and corpus callosum demonstrated excellent correlation with manually-corrected volumes. In addition, group comparison statistical inferences (i.e., whether or not volume differences between former NFL players and controls reached statistical significance) were the same whether using automated segmentation or manually-corrected lateral ventricle volumes and corpus callosum volumes. Similarly, statistical inferences when correlating former NFL player volumes with neurobehavioral scores were the same whether using automated segmentation or manually-corrected lateral ventricle volumes and corpus callosum volumes.

The between-group mean volume differences measured with and without manual correction are very similar for the corpus callosum. For the lateral ventricles, there is a slightly larger difference in the mean volumes when measured with and without manual correction. However, the ranges of values are quite large and the difference remains statistically insignificant. Moreover, the magnitude of the mean volume differences between former NFL players and controls is driven in part by the mixed effects model's correlation between the hemispheres and the model's incorporation of and control for the confounders of intracranial volume and age. For example, the absolute raw difference in size of the left lateral ventricle between former NFL players and controls is 1663 mL with manually corrected volumes and 736 mL (0.02%) with automatically segmented volumes, compared to 2300 mL and 1139 mL as calculated with the mixed effects model. The trend toward slightly higher differences measured with manual correction is described below in Section 4.3.

Post-mortem pathology studies have identified corpus callosum abnormalities in multiple diseases with neurodegenerative and psychiatric components including in schizophrenia (Bigelow et al., 1983;

Rosenthal and Bigelow, 1972), multiple sclerosis (Evangelou et al., 2000), traumatic brain injury (Anderson and Bigler, 1994), and chronic traumatic encephalopathy (McKee et al., 2013; Mez et al., 2017). Neuroimaging studies have identified corpus callosum abnormalities in-vivo in both traumatic brain injury (see reviews by Shenton et al., 2012 and Mu et al., 2017) and in repetitive head impacts (see reviews by Ng et al., 2014 and Koerte et al., 2015), both of which are sustained by American football players. Similarly, abnormal enlargement of the lateral ventricles has been described in chronic traumatic encephalopathy (McKee et al., 2013; Mez et al., 2017) and in other diseases that have neurodegenerative features including multiple sclerosis (e.g., Turner et al., 2003), schizophrenia (e.g., Kempton et al., 2010), Alzheimer's disease (e.g., Nestor et al., 2008), and alcoholism (e.g., Fox et al., 1976). Both corpus callosum volume and lateral ventricle volume could therefore potentially serve as in-vivo biomarkers for neurodegenerative processes.

Given the excellent correlation of automated corpus callosum and lateral ventricle segmentation volumes generated by FreeSurfer with manually-corrected volumes, and the concordance of results when using either automated or manually-corrected volumes in this study, the corpus callosum and lateral ventricles can probably be reliably segmented by FreeSurfer without the need for manual editing, at least in large-scale studies.

4.2. Amygdala, hippocampus, and cingulate gyrus

In this study, automated segmentation volumes of the amygdala, hippocampus, and cingulate gyrus demonstrated poorer correlation with manually-corrected volumes than was the case for the corpus callosum and the lateral ventricles. Furthermore, statistical inferences differed when comparing volume differences between former NFL players and controls and also when evaluating former NFL player volume correlations with neurobehavioral scores, depending on whether automated segmentation or manually-corrected amygdala, hippocampus, and cingulate gyrus volumes were used. These findings are corroborated by markedly discrepant mean volume differences and estimated neurobehavioral score effect sizes in these brain regions, particularly in the cingulate cortices, when using automated segmentation as opposed to manually-corrected volumes.

The correlations of automated segmentation and manually-corrected volumes of the amygdala and hippocampus were however, somewhat improved when combined into a single amygdala-hippocampal complex volume. Moreover, the automated segmentation and manually-corrected amygdala-hippocampal complex volumes yielded the same statistical study results when comparing the volumes of former NFL players and control subjects and yielded 7 concordant results out of 8 when evaluating the relationships of former NFL player volumes with neurobehavioral scores. Similarly, the discrepancies between the mean volume differences and the estimated neurobehavioral score effect sizes are smaller than those seen when evaluating the amygdala and hippocampus as separate structures.

The amygdala, hippocampus, and cingulate gyrus are of great interest in neuroimaging studies of repetitive head impacts (see reviews by Ng et al., 2014 and Koerte et al., 2015), and mild traumatic brain injury (see reviews by Shenton et al., 2012 and Mu et al., 2017). In chronic traumatic encephalopathy, post-mortem studies have identified deposition of hyperphosphorylated tau in the limbic system, particularly in the amygdala and hippocampus but also to a lesser degree in the cingulate gyrus (McKee et al., 2013; Mez et al., 2017; Omalu et al., 2011). As with the corpus callosum and the lateral ventricle volumes, the volumes of the amygdala, hippocampus, and cingulate gyrus could thus potentially serve as in-vivo biomarkers for diseases with neurodegenerative features.

Mirroring the results of several other studies that demonstrated correlation coefficients often substantially < 0.8, this study demonstrates suboptimal correlation of automated segmentation volumes with manually-corrected volumes of the amygdala and hippocampus (Cherbuin et al., 2009; de Flores et al., 2015; Grimm et al., 2015; Morey et al., 2009; Schoemaker et al., 2016; Wenger et al., 2014). This study also demonstrates that, at least in this cohort of former professional football players with history of exposure to repetitive head impacts, failing to manually edit the amygdala and hippocampus volumes generated by FreeSurfer led to very different brain volume group comparison and neurobehavioral correlation study results. Automated segmentation volumes of the amygdala, hippocampus, and cingulate gyrus are thus not adequate for meaningful study on their own without further manual editing.

Results are substantially improved, however, when the amygdala and hippocampus volumes are combined into an amygdala-hippocampal complex. The combined amygdala-hippocampal complex can thus possibly be segmented by FreeSurfer without the need for manual editing for large-scale studies, provided appropriate quality control is performed.

4.3. Technical considerations

There was greater variance in the sizes of the studied brain structures when the volumes were manually-edited, suggesting that the method of segmentation employed by FreeSurfer, which uses an atlas based approach on a training set to label the likely structural location of each voxel (Fischl et al., 2004), may artificially improve precision. This artificial precision may be related to the inherently normalizing process of stretching a fixed image atlas to brains that differ from the atlas in unique and variable ways that are better captured in the manual correction process. The results of a prior study, which showed that FreeSurfer relatively overestimates the size of smaller hippocampi but not larger hippocampi (Wenger et al., 2014), supports this hypothesis. Although artificial precision could be expected to spuriously increase statistical power, this study demonstrates more statistically significant results with manually-edited volumes, a finding presumably due to better accuracy achieved through the manual editing process because the manual editing process allows for careful evaluation of deep brain structure with low-contrast borders that may not be within the resolving power of FreeSurfer. In addition, as previously described by Wenger et al., 2014, manual segmentation and manual editing typically follow rules that define the often unclear borders between amygdala and the hippocampus and between the tail of the hippocampus and the lateral ventricle whereas FreeSurfer tends to be more inclusive in these areas (Fig. 1). Although Wenger et al. used a set of rules described by Pruessner et al. (2000) and this study used a set of rules described by Gurvits et al. (1996), this study replicates the qualitative results described by Wenger et al.

4.4. Limitations

Just as there are significant differences between the results obtained with FreeSurfer 5 and prior versions of FreeSurfer (Gronenschild et al., 2012), the results of this paper may therefore not apply to other versions of FreeSurfer. Additionally, this study evaluated differences between measured volumes, but did not account for volume overlap or shape and the results may not be translatable to volume overlap or shape (Morey et al., 2009). The DETECT study also included only male subjects and the results of this study may therefore not be generalizable to female subjects. The DETECT study subjects were former professional football players exposed to repetitive head impacts and the conclusions drawn regarding the effect of using automated segmentation versus manually-corrected volumes on group comparison and neurobehavioral study results may not be translatable to studies of subjects with different pathologies. However, the conclusions drawn regarding automated segmentation and manually-corrected volume correlations are likely generalizable. Although the data set for this study appears too small to benefit from the multiple imputation statistical analysis

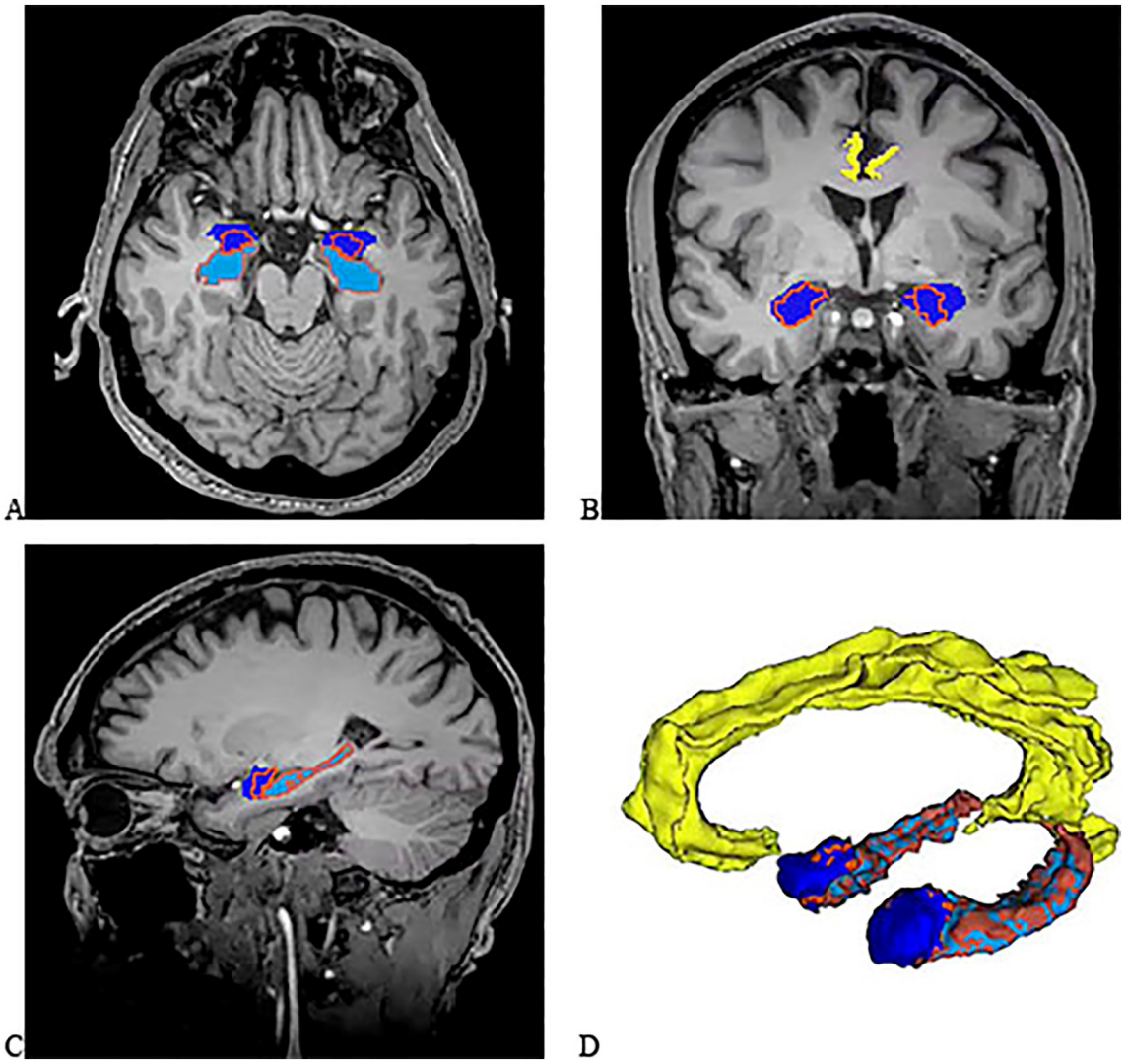


Fig. 1. (A) Axial 2D, (B) Coronal 2D, and (C) Sagittal 2D T1-weighted images of the brain. The manually-edited label map surfaces of the right and left amygdala are dark blue while the manually-edited label map surfaces of the right and left hippocampus are light blue. The outline of the label maps as drawn by FreeSurfer without manual correction are superimposed as orange outlines. The cingulate gyrus label map is presented in yellow. (D) Obliqued 3D reconstruction of the FreeSurfer generated label maps (orange) and manually-edited label maps (blue) created from a randomly selected DETECT study subject using the Model Maker module of Slicer 4.5.0. Note the manually-edited volume of the amygdala is substantially larger than the unedited volume as segmented by FreeSurfer. Also note that the unedited FreeSurfer volume of the hippocampus extends more posterior than the manually-edited volume. The label map was not modified for the 2D image creation but was smoothed for the 3D reconstruction using a sinc filter with a smoothing factor of 10. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

method, which can reduce the number of label maps that require manual correction (Chua et al., 2015), studies with larger data sets may benefit from that approach. The image processing pipeline for this study did not utilize the -T2pial or -FLAIRpial options to optimize pial surface estimations. However, these options are typically useful when there is dura within the brain mask that is not adequately removed with skull stripping. In the utilized pipeline, each brain mask was reviewed and manually edited if necessary, thus likely yielding more precise results than the available automated FreeSurfer options.

5. Conclusions

Automated segmentation using FreeSurfer 5.3 yields excellent correlation with manually-edited volumes of the corpus callosum and lateral ventricles but suboptimal correlation for amygdala, hippocampus, and cingulate gyrus. In addition, using automated segmentation volumes leads to substantially different study results than using manually-corrected volumes when correlating brain volumes with neurobehavioral test scores in this cohort of former professional football players. Study result concordance is improved when the amygdala and hippocampus volumes are combined into an amygdala-hippocampal complex. Automated FreeSurfer-derived segmentation volumes of the corpus callosum and lateral ventricles, and amygdala-hippocampus complex may therefore be suitable for analysis without manual correction, provided appropriate quality control is performed. However, automated FreeSurfer-derived segmentation volumes of the amygdala, hippocampus, and cingulate gyrus should not be utilized for analysis without manual correction until further refinements are made to the FreeSurfer algorithm and appropriately tested.

Author disclosure statement

Dr. Robert Stern is a paid consultant to Avanir Pharmaceuticals, Inc. (Aliso Viejo, CA), Biogen (Cambridge, MA), and Eli Lilly (Indianapolis, IN). He receives royalties for published neuropsychological tests from Psychological Assessment Resources, Inc. (Lutz, FL, USA). He is on the Board of Directors of King-Devick Technologies, Inc. (Chicago, IL), and is a member of the NFL Players Association Mackey-White Health and

Safety Committee. For all other authors, no competing financial interests exist.

Acknowledgments

This study was supported by the NIH (R01 NS 078337 (RAS, YT, CMB, NGF, BMM, CC); R01 NS100952 (IK, APL); F31 NS 081957 (JMS); 1F32NS096803-01 (MLA), P30 AG13846 (RAS, YT, MLA, CMB, BMM, CC); P41 EB015902 (OP); T32GM074905 (IW)), and participant travel was partially funded by JetBlue Airlines, the National Football League, and the NFL Players Association. This study was also partly supported by the Else Kröner-Fresenius Foundation, Germany (IK), by LMU Munich's Institutional Strategy LMUexcellent within the framework of the German Excellence Initiative, and by a VA Merit Award (MES, MC). PW and VS were supported by the German Academic Exchange Service PROMOS award. PW was also supported by an international research scholarship by the medical faculty of the Ludwig-Maximillian University Munich. This work was completed in partial fulfillment of VS's dissertation. CL was supported by the Canadian Institutes of Health Research Frederick Banting and Charles Best Doctoral Award. The sponsors had no role in the design and conduct of the study, the collection, analysis, or the interpretation of the data, nor in the preparation, review, or approval of the manuscript, or the decision of submission for publication.

References

Ahmed-Leitao, F., Spies, G., van den Heuvel, L., Seedat, S., 2016. Hippocampal and amygdala volumes in adults with posttraumatic stress disorder secondary to childhood abuse or maltreatment: a systematic review. Psychiatry Res. 256, 33–43. http://dx.doi.org/10.1016/j.pscychresns.2016.09.008.

Akhondi-Asl, A., Jafari-Khouzani, K., Elisevich, K., Soltanian-Zadeh, H., 2011. Hippocampal volumetry for lateralization of temporal lobe epilepsy: automated versus manual methods. NeuroImage 54 (Suppl. 1), S218–226. http://dx.doi.org/10.1016/j.neuroimage.2010.03.066.

Alosco, M.L., Jarnagin, J., Tripodis, Y., Platt, M., Martin, B., Chaisson, C.E., Baugh, C.M., Fritts, N.G., Cantu, R.C., Stern, R.A., 2016. Olfactory function and associated clinical correlates in former national football league players. J. Neurotrauma. http://dx.doi.org/10.1089/neu.2016.4536.

Alosco, M.L., Tripodis, Y., Jarnagin, J., Baugh, C.M., Martin, B., Chaisson, C.E., Estochen, N., Song, L., Cantu, R.C., Jeromin, A., Stern, R.A., 2017. Repetitive head impact exposure and later-life plasma total tau in former national football league players. Alzheimers Dement. Amst. Neth. 7, 33–40. http://dx.doi.org/10.1016/j.dadm.2016.11.003.

Anderson, C.V., Bigler, E.D., 1994. The role of caudate nucleus and corpus callosum atrophy in trauma-induced anterior horn dilation. Brain Inj. 8, 565–569. http://dx.doi.org/10.3109/02699059409151008.

Barron, S.A., Jacobs, L., Kinkel, W.R., 1976. Changes in size of normal lateral ventricles during aging determined by computerized tomography. Neurology 26, 1011–1013.

Bernick, C., Banks, S.J., Shin, W., Obuchowski, N., Butler, S., Noback, M., Phillips, M., Lowe, M., Jones, S., Modic, M., 2015. Repeated head trauma is associated with smaller thalamic volumes and slower processing speed: the professional fighters' brain health study. Br. J. Sports Med. 49, 1007–1011. http://dx.doi.org/10.1136/bjsports-2014-093877.

Bigelow, L.B., Nasrallah, H.A., Rauscher, F.P., 1983. Corpus callosum thickness in chronic schizophrenia. Br. J. Psychiatry 142, 284–287. http://dx.doi.org/10.1192/bjp.142.3.284.

Buckner, R.L., Head, D., Parker, J., Fotenos, A.F., Marcus, D., Morris, J.C., Snyder, A.Z., 2004. A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. NeuroImage 23, 724–738. http://dx.doi.org/10.1016/j.neuroimage.2004.06.018.

Busatto, G.F., Diniz, B.S., Zanetti, M.V., 2008. Voxel-based morphometry in Alzheimer's disease. Expert. Rev. Neurother. 8, 1691–1702. http://dx.doi.org/10.1586/14737175.8.11.1691.

Chera, B.S., Amdur, R.J., Patel, P., Mendenhall, W.M., 2009. A radiation oncologist's guide to contouring the hippocampus. Am. J. Clin. Oncol. 32, 20–22. http://dx.doi.org/10.1097/COC.0b013e318178e4e8.

Cherbuin, N., Anstey, K.J., Réglade-Meslin, C., Sachdev, P.S., 2009. In vivo hippocampal measurement and memory: a comparison of manual tracing and automated segmentation in a large community-based sample. PLoS One 4, e5265. http://dx.doi.org/10.1371/journal.pone.0005265.

Chua, A.S., Egorova, S., Anderson, M.C., Polgar-Turcsanyi, M., Chitnis, T., Weiner, H.L., Guttmann, C.R.G., Bakshi, R., Healy, B.C., 2015. Using multiple imputation to efficiently correct cerebral MRI whole brain lesion and atrophy data in patients with multiple sclerosis. NeuroImage 119, 81–88. http://dx.doi.org/10.1016/j.neuroimage.2015.06.037.

Cicchetti, D.V., 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instrument in psychology. Psychol. Assess. 6, 284–290.

Crisco, J.J., Wilcox, B.J., Beckwith, J.G., Chu, J.J., Duhaime, A.-C., Rowson, S., Duma, S.M., Maerlender, A.C., McAllister, T.W., Greenwald, R.M., 2011. Head impact exposure in collegiate football players. J. Biomech. 44, 2673–2678. http://dx.doi.org/10.1016/j.jbiomech.2011.08.003.

Dill, V., Franco, A.R., Pinho, M.S., 2015. Automated methods for hippocampus segmentation: the evolution and a review of the state of the art. Neuroinformatics 13, 133–150. http://dx.doi.org/10.1007/s12021-014-9243-4.

Evangelou, N., Konz, D., Esiri, M.M., Smith, S., Palace, J., Matthews, P.M., 2000. Regional axonal loss in the corpus callosum correlates with cerebral white matter lesion volume and distribution in multiple sclerosis. Brain 123, 1845–1849. http://dx.doi.org/10.1093/brain/123.9.1845.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J.V., Pieper, S., Kikinis, R., 2012. 3D slicer as an image computing platform for the quantitative imaging network. Magn. Reson. Imaging 30, 1323–1341. http://dx.doi.org/10.1016/j.mri.2012.05.001.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex N. Y. N. 1991 (14), 11–22.

de Flores, R., La Joie, R., Landeau, B., Perrotin, A., Mézenge, F., de La Sayette, V., Eustache, F., Desgranges, B., Chételat, G., 2015. Effects of age and Alzheimer's disease on hippocampal subfields: comparison between manual and FreeSurfer volumetry. Hum. Brain Mapp. 36, 463–474. http://dx.doi.org/10.1002/hbm.22640.

Fox, J.H., Ramsey, R.G., Huckman, M.S., Proske, A.E., 1976. Cerebral ventricular enlargement. Chronic alcoholics examined by computerized tomography. JAMA 236, 365–368.

Goddeyne, C., Nichols, J., Wu, C., Anderson, T., 2015. Repetitive mild traumatic brain injury induces ventriculomegaly and cortical thinning in juvenile rats. J. Neurophysiol. 113, 3268–3280. http://dx.doi.org/10.1152/jn.00970.2014.

González-Villà, S., Oliver, A., Valverde, S., Wang, L., Zwiggelaar, R., Lladó, X., 2016. A review on brain structures segmentation in magnetic resonance imaging. Artif. Intell. Med. 73, 45–69. http://dx.doi.org/10.1016/j.artmed.2016.09.001.

Grimm, O., Pohlack, S., Cacciaglia, R., Winkelmann, T., Plichta, M.M., Demirakca, T., Flor, H., 2015. Amygdalar and hippocampal volume: a comparison between manual segmentation, Freesurfer and VBM. J. Neurosci. Methods 253, 254–261. http://dx.doi.org/10.1016/j.jneumeth.2015.05.024.

Gronenschild, E.H.B.M., Habets, P., Jacobs, H.I.L., Mengelers, R., Rozendaal, N., van Os, J., Marcelis, M., 2012. The effects of FreeSurfer version, workstation type, and Macintosh operating system version on anatomical volume and cortical thickness measurements. PLoS One 7, e38234. http://dx.doi.org/10.1371/journal.pone.0038234.

Gurvits, T.V., Shenton, M.E., Hokama, H., Ohta, H., Lasko, N.B., Gilbertson, M.W., Orr, S.P., Kikinis, R., Jolesz, F.A., McCarley, R.W., Pitman, R.K., 1996. Magnetic resonance imaging study of hippocampal volume in chronic, combat-related posttraumatic stress disorder. Biol. Psychiatry 40, 1091–1099. http://dx.doi.org/10.1016/S0006-3223(96)00229-6.

Haller, S., Zaharchuk, G., Thomas, D.L., Lovblad, K.-O., Barkhof, F., Golay, X., 2016. Arterial spin labeling perfusion of the brain: emerging clinical applications. Radiology 281, 337–356. http://dx.doi.org/10.1148/radiol.2016150789.

Hulshoff Pol, H.E., Kahn, R.S., 2008. What happens after the first episode? A review of progressive brain changes in chronically ill patients with schizophrenia. Schizophr. Bull. 34, 354–366. http://dx.doi.org/10.1093/schbul/sbm168.

Inglese, P., Amoroso, N., Boccardi, M., Bocchetta, M., Bruno, S., Chincarini, A., Errico, R., Frisoni, G.B., Maglietta, R., Redolfi, A., Sensi, F., Tangaro, S., Tateo, A., Bellotti, R., Alzheimer's Disease Neuroimaging Initiative, 2015. Multiple RF classifier for the hippocampus segmentation: method and validation on EADC-ADNI harmonized hippocampal protocol. Phys. Medica PM Int. J. Devoted Appl. Phys. Med. Biol. Off. J. Ital. Assoc. Biomed. Phys. AIFB 31, 1085–1091. http://dx.doi.org/10.1016/j.ejmp.2015.08.003.

Kantarci, K., Jack, C.R., 2003. Neuroimaging in Alzheimer disease: an evidence-based review. Neuroimaging Clin. N. Am. 13, 197–209.

Kempton, M.J., Stahl, D., Williams, S.C.R., DeLisi, L.E., 2010. Progressive lateral ventricular enlargement in schizophrenia: a meta-analysis of longitudinal MRI studies. Schizophr. Res. 120, 54–62. http://dx.doi.org/10.1016/j.schres.2010.03.036.

Kiernan, J.A., 2012. Anatomy of the temporal lobe. Epilepsy Res. Treat. 2012, 176157. http://dx.doi.org/10.1155/2012/176157.

Koerte, I.K., Lin, A.P., Willems, A., Muehlmann, M., Hufschmidt, J., Coleman, M.J., Green, I., Liao, H., Tate, D.F., Wilde, E.A., Pasternak, O., Bouix, S., Rathi, Y., Bigler, E.D., Stern, R.A., Shenton, M.E., 2015. A review of neuroimaging findings in repetitive brain trauma. Brain Pathol. Zurich Switz. 25, 318–349. http://dx.doi.org/10.1111/bpa.12249.

Laurent, S., Thibaud, J.L., Hordeaux, J., Reyes-Gomez, E., Delisle, F., Blot, S., Colle, M.A., 2010. Chronic traumatic brain injury in a dog. J. Comp. Pathol. 143, 75–80. http://dx.doi.org/10.1016/j.jcpa.2009.12.008.

Makris, N., Gasic, G.P., Kennedy, D.N., Hodge, S.M., Kaiser, J.R., Lee, M.J., Kim, B.W., Blood, A.J., Evins, A.E., Seidman, L.J., Iosifescu, D.V., Lee, S., Baxter, C., Perlis, R.H., Smoller, J.W., Fava, M., Breiter, H.C., 2008. Cortical thickness abnormalities in cocaine addiction—a reflection of both drug use and a pre-existing disposition to drug abuse? Neuron 60, 174–188. http://dx.doi.org/10.1016/j.neuron.2008.08.011.

McKee, A.C., Cantu, R.C., Nowinski, C.J., Hedley-Whyte, E.T., Gavett, B.E., Budson, A.E., Santini, V.E., Lee, H.-S., Kubilus, C.A., Stern, R.A., 2009. Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury. J. Neuropathol. Exp. Neurol. 68, 709–735. http://dx.doi.org/10.1097/NEN.

0b013e3181a9d503.
McKee, A.C., Stern, R.A., Nowinski, C.J., Stein, T.D., Alvarez, V.E., Daneshvar, D.H., Lee, H.-S., Wojtowicz, S.M., Hall, G., Baugh, C.M., Riley, D.O., Kubilus, C.A., Cormier, K.A., Jacobs, M.A., Martin, B.R., Abraham, C.R., Ikezu, T., Reichard, R.R., Wolozin, B.L., Budson, A.E., Goldstein, L.E., Kowall, N.W., Cantu, R.C., 2013. The spectrum of disease in chronic traumatic encephalopathy. Brain J. Neurol. 136, 43–64. http://dx.doi.org/10.1093/brain/aws307.
Mendrik, A.M., Vincken, K.L., Kuijf, H.J., Breeuwer, M., Bouvy, W.H., de Bresser, J., Alansary, A., de Bruijne, M., Carass, A., El-Baz, A., Jog, A., Katyal, R., Khan, A.R., van der Lijn, F., Mahmood, Q., Mukherjee, R., van Opbroek, A., Paneri, S., Pereira, S., Persson, M., Rajchl, M., Sarikaya, D., Smedby, Ö., Silva, C.A., Vrooman, H.A., Vyas, S., Wang, C., Zhao, L., Biessels, G.J., Viergever, M.A., 2015. MRBrainS challenge: online evaluation framework for brain image segmentation in 3T MRI scans. Comput. Intell. Neurosci. 2015, 813696. http://dx.doi.org/10.1155/2015/813696.
Mez, J., Daneshvar, D.H., Kiernan, P.T., Abdolmohammadi, B., Alvarez, V.E., Huber, B.R., Alosco, M.L., Solomon, T.M., Nowinski, C.J., McHale, L., Cormier, K.A., Kubilus, C.A., Martin, B.M., Murphy, L., Baugh, C.M., Montenigro, P.H., Chaisson, C.E., Tripodis, Y., Kowall, N.W., Weuve, J., McClean, M.D., Cantu, R.C., Goldstein, L.E., Katz, D.I., Stern, R.A., Stein, T.D., McKee, A.C., 2017. Clinicopathological evaluation of chronic traumatic encephalopathy in players of American football. JAMA 318, 360–370. http://dx.doi.org/10.1001/jama.2017.8334.
Milani, A.C.C., Hoffmann, E.V., Fossaluza, V., Jackowski, A.P., Mello, M.F., 2017. Does pediatric post-traumatic stress disorder alter the brain? Systematic review and meta-analysis of structural and functional magnetic resonance imaging studies. Psychiatry Clin. Neurosci. 71, 154–169. http://dx.doi.org/10.1111/pcn.12473.
Morey, R.A., Petty, C.M., Xu, Y., Hayes, J.P., Wagner, H.R., Lewis, D.V., LaBar, K.S., Styner, M., McCarthy, G., 2009. A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. NeuroImage 45, 855–866. http://dx.doi.org/10.1016/j.neuroimage.2008.12.033.
Mu, W., Catenaccio, E., Lipton, M.L., 2017. Neuroimaging in blast-related mild traumatic brain injury. J. Head Trauma Rehabil. 32, 55–69. http://dx.doi.org/10.1097/HTR.0000000000000213.
Næss-Schmidt, E., Tietze, A., Blicher, J.U., Petersen, M., Mikkelsen, I.K., Coupé, P., Manjón, J.V., Eskildsen, S.F., 2016. Automatic thalamus and hippocampus segmentation from MP2RAGE: comparison of publicly available methods and implications for DTI quantification. Int. J. Comput. Assist. Radiol. Surg. http://dx.doi.org/10.1007/s11548-016-1433-0.
Nestor, S.M., Rupsingh, R., Borrie, M., Smith, M., Accomazzi, V., Wells, J.L., Fogarty, J., Bartha, R., Alzheimer's Disease Neuroimaging Initiative, 2008. Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. Brain J. Neurol. 131, 2443–2454. http://dx.doi.org/10.1093/brain/awn146.
Ng, T.S., Lin, A.P., Koerte, I.K., Pasternak, O., Liao, H., Merugumala, S., Bouix, S., Shenton, M.E., 2014. Neuroimaging in repetitive brain trauma. Alzheimers Res. Ther. 6, 10. http://dx.doi.org/10.1186/alzrt239.
Olabi, B., Ellison-Wright, I., McIntosh, A.M., Wood, S.J., Bullmore, E., Lawrie, S.M., 2011. Are there progressive brain changes in schizophrenia? A meta-analysis of structural magnetic resonance imaging studies. Biol. Psychiatry 70, 88–96. http://dx.doi.org/10.1016/j.biopsych.2011.01.032.
Omalu, B., Bailes, J., Hamilton, R.L., Kamboh, M.I., Hammers, J., Case, M., Fitzsimmons, R., 2011. Emerging histomorphologic phenotypes of chronic traumatic encephalopathy in American athletes. Neurosurgery 69, 173–183. discussion 183. https://doi.org/10.1227/NEU.0b013e318212bc7b.
Pruessner, J.C., Li, L.M., Serles, W., Pruessner, M., Collins, D.L., Kabani, N., Lupien, S., Evans, A.C., 2000. Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. Cereb. Cortex N. Y. N 1991 (10), 433–442.
Rosenthal, R., Bigelow, L.B., 1972. Quantitative brain measurements in chronic schizophrenia. Br. J. Psychiatry J. Ment. Sci. 121, 259–264.
Schoemaker, D., Buss, C., Head, K., Sandman, C.A., Davis, E.P., Chakravarty, M.M., Gauthier, S., Pruessner, J.C., 2016. Hippocampus and amygdala volumes from magnetic resonance images in children: assessing accuracy of FreeSurfer and FSL against manual segmentation. NeuroImage 129, 1–14. http://dx.doi.org/10.1016/j.neuroimage.2016.01.038.
Shenton, M.E., Whitford, T.J., Kubicki, M., 2010. Structural neuroimaging in schizophrenia: from methods to insights to treatments. Dialogues Clin. Neurosci. 12, 317–332.
Shenton, M.E., Hamoda, H.M., Schneiderman, J.S., Bouix, S., Pasternak, O., Rathi, Y., Vu, M.-A., Purohit, M.P., Helmer, K., Koerte, I., Lin, A.P., Westin, C.-F., Kikinis, R., Kubicki, M., Stern, R.A., Zafonte, R., 2012. A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury. Brain Imaging Behav. 6, 137–192. http://dx.doi.org/10.1007/s11682-012-9156-5.
Stamm, J.M., Koerte, I.K., Muehlmann, M., Pasternak, O., Bourlas, A.P., Baugh, C.M., Giwerc, M.Y., Zhu, A., Coleman, M.J., Bouix, S., Fritts, N.G., Martin, B.M., Chaisson, C., McClean, M.D., Lin, A.P., Cantu, R.C., Tripodis, Y., Stern, R.A., Shenton, M.E., 2015. Age at first exposure to football is associated with altered corpus callosum white matter microstructure in former professional football players. J. Neurotrauma 32, 1768–1776. http://dx.doi.org/10.1089/neu.2014.3822.
Stern, R.A., Tripodis, Y., Baugh, C.M., Fritts, N.G., Martin, B.M., Chaisson, C., Cantu, R.C., Joyce, J.A., Shah, S., Ikezu, T., Zhang, J., Gercel-Taylor, C., Taylor, D.D., 2016. Preliminary study of plasma exosomal tau as a potential biomarker for chronic traumatic encephalopathy. J. Alzheimers Dis. JAD 51, 1099–1109. http://dx.doi.org/10.3233/JAD-151028.
Turner, B., Lin, X., Calmon, G., Roberts, N., Blumhardt, L.D., 2003. Cerebral atrophy and disability in relapsing-remitting and secondary progressive multiple sclerosis over four years. Mult. Scler. Houndmills Basingstoke Engl. 9, 21–27.
Wenger, E., Mårtensson, J., Noack, H., Bodammer, N.C., Kühn, S., Schaefer, S., Heinze, H.-J., Düzel, E., Bäckman, L., Lindenberger, U., Lövdén, M., 2014. Comparing manual and automatic segmentation of hippocampal volumes: reliability and validity issues in younger and older brains. Hum. Brain Mapp. 35, 4236–4248. http://dx.doi.org/10.1002/hbm.22473.
Wible, C.G., Shenton, M.E., Fischer, I.A., Allard, J.E., Kikinis, R., Jolesz, F.A., Iosifescu, D.V., McCarley, R.W., 1997. Parcellation of the human prefrontal cortex using MRI. Psychiatry Res. 76, 29–40.

5. PUBLICATION B

5.1 Background

The original article B with the title "Age at First Exposure to Repetitive Head Impacts Is Associated with Smaller Thalamic Volumes in Former Professional American Football Players" was published in January 2018 in the *Journal of Neurotrauma* (IF: 3.75, citations: 31). In this work we aimed to investigate the association between AFE to American Football and thalamic volumes in a sample of symptomatic former professional football players at risk of CTE. We also investigated the association between total years of playing football and thalamic volumes. To evaluate the results in a clinical context, we also explored the associations between thalamic volumes and neurobehavioral measures.

5.2 Methods

The same eighty-six symptomatic former professional football players (mean age = 54.9 ± 7.9 years) from publication A were included in this analysis. Thalamic volumes were derived using FreeSurfer 5.3 for every participant. Mood and behavior, attention/psychomotor speed, and visual and verbal memory were assessed in a comprehensive neurobehavioral testing battery. Further, measures of exposure, including total years of playing football and AFE of playing football, were compiled. All statistical analyses were performed with Statistical Analysis System (SAS version 9.4; SAS Institute Inc., North Carolina, USA). We investigated the association between thalamic volumes and total years of play and AFE as well as neurobehavioral test results using mixed effects regression models. Written informed consent was obtained from all participants prior to enrollment.

5.3 Results

The associations between thalamic volume and measures of exposure were statistically significant. Right and left thalamic volume negatively correlated with total years of play (left, $p = 0.012$; right, $p = 0.03$) indicating that the longer athletes played football, the smaller their thalamic volumes. In fact, the right thalamus was on average 38.6 mm^3 and the left thalamus 53.2 mm^3 smaller for every year of playing football. Further, right thalamic volume was associated with AFE, even when adjusting for total years of play ($p = 0.014$). This means that the younger the athlete's age when he started playing football, the smaller his right thalamus. Importantly, the impact of AFE on right thalamic volume was almost twice the effect of total years of play with an average decrease of 64.9 mm^3 for every year an athlete started playing earlier compared to a decrease of 38.6 mm^3 for every year an athlete played longer. Moreover, there were statistically significant positive associations between left thalamic volume and visual memory ($r = 0.28$, $p = 0.014$) as well as right thalamic volume and mood and behavior ($r = 0.34$, $p = 0.003$). The smaller the left thalamic volume, the worse the performance in visual memory and the larger the right thalamic volume, the worse the mood and behavioral symptoms.

5.4 Discussion

Publication B investigated the association between thalamic volumes and measures of exposure, in particular AFE to American Football, and in a second step explored the association between thalamic volumes with neurobehavioral parameters.

The results showed that the more years athletes participated in football, the smaller their left and right thalamic volumes. A decrease in thalamic volumes has previously been reported in neuropathologically confirmed cases of CTE.[2] Further, the results of this study are in line with findings in living athletes after concussion and exposure to RHI.[36,156] The results are also in line with the study by Bernick et al. among active professional fighters.[36,37] Although head accelerations sustained during professional boxing exceed those sustained during professional American Football[44,45] we saw a similar association between the total impact burden and thalamic volumes.

This is the first study showing that AFE to RHI in football affects subcortical gray matter volume. In our study younger AFE to football was related with smaller right thalamic volumes even after adjusting for total years of play. This is noteworthy because an earlier age for starting to play football could likely be linked to more years of play and thus have driven the results. In fact, the effect of AFE on right thalamic volume was almost double the effect of total years of play with a decrease of 64.9 mm^3 for every year athletes had started playing earlier and a decrease of 38.6 mm^3 for every year athletes had played longer. The finding is consistent with three previous studies on AFE in living athletes.[50,64-65] They showed that younger AFE was associated with microstructural alterations in the corpus callosum and also linked to more severe cognitive and mood and behavioral impairment later in life.[50,64-65] Our results suggest that RHI may have an especially strong impact on the developing brain potentially due to interference with critical periods of brain development and maturation as further discussed in publication B.

Findings from this study also showed that symptoms associated with CTE may be related to thalamic volume. The thalamus is involved in a great number of essential brain functions and is also an important relay station for the cortex.[157] Among the processes the thalamus is involved in are memory functions, the circadian rhythm, emotions, sensorimotor control, as well as processing of multiple forms of sensory input (such as taste, visual, and auditory information) before transmitting it to the cortex.[157] In fact, in this study we showed that a larger right thalamic volume was related to worse mood and behavioral symptoms. This result is consistent with a finding among medication-naïve depressive patients and indicates that the players with larger thalamic volumes may not have undergone medical treatment for mood and behavioral symptoms.[158] In this study we also found that smaller left thalamic volume was associated with worse visual memory performance which is in line with a previous study on severe traumatic brain injury.[159] In fact, deficits in memory function are among the most commonly investigated dysfunctions after thalamic injury.[157] A delay in visual memory tasks is one of the most common findings in thalamic damage and has, for example, often been described in patients with ischemic thalamic stroke.[157]

Among the reasons for a thalamic volume decrease may be processes associated with RHI, such as microgliosis.[160] As of today, brain global and regional atrophy as a result of neurodegenerative processes are not reversible. Therefore, treatment goal should be to maintain the status quo and prolong the progress of cognitive decline, e.g. by modifying environmental factors or offering psychotherapy.[157]

As mentioned previously, the core goal in CTE research to date is to find a diagnostic marker *in vivo*. However, despite thalamic volume reduction has been reported in this sample at risk of CTE, several reasons exist why a

decrease in thalamic volume may not serve as a marker for early diagnosis of CTE in the living. First, although thalamic atrophy has been described in neuropathologically confirmed cases of CTE, it is not specific for CTE and has been described in other neuropathologies, such as schizophrenia[161], multiple sclerosis[162], or substance use disorder[163]. Moreover, thalamic atrophy was reported in advanced stages of neuropathologically confirmed cases of CTE, making it per se less attractive for an imaging marker. The goal would be to enable a diagnosis in an early stage of the disease with almost no or few symptoms due to hopefully established treatment options in the future. However, even at this point where there is no specific treatment an early diagnosis would enable physicians and patients to help modify environmental factors, offer psychotherapy, and elaborate therapies to maintain the status quo without or slower further decline in cognitive or mood and motor functions. Second, volumetric analyses using FreeSurfer are not established in clinical settings today. Future research should further investigate if the combined analysis of certain anatomical structures often shown affected in CTE is a promising option. Importantly, PET tau-imaging has shown encouraging results.[164] Ultimately, although the participants we included in this study certainly are at risk of CTE, without *postmortem* analysis we cannot be sure if they really suffer from CTE or if their symptoms are a manifestation of different forms of pathology.
Moreover, this study was conducted among an all-male sample. Future research should focus also on females given their rising number in sports and the initial evidence that the female brain may be more vulnerable to RHI compared with their male counterparts.[165] Ultimately, prospective studies with a longitudinal study design should be conducted to better understand associations between exposure to RHI and related brain structural and functional changes. Longitudinal studies are also urgently needed to further elucidate the effect on the immature brain in order to implement adjustments in youth football policies and regulations accordingly. In this study the cross-sectional study design precluded a causal link between RHI and thalamic volume as well as AFE and thalamic volume. Further, without *postmortem* analysis it is not possible to determine whether or not our findings are specifically related to CTE or to other long-term consequences of RHI.

5.5 Contribution

During my research year at the Psychiatry Neuroimaging Laboratory in Boston, USA, I worked at this unique and state-of-the-art dataset. My contribution to this study consisted of the following:

1. Data analysis and interpretation.
2. Writing the manuscript.
3. Presenting the results orally at IBIA World Congress on Brain Injury in New Orleans, USA, in 2017.

JOURNAL OF NEUROTRAUMA 35:278–285 (January 15, 2018)

DOI: 10.1089/neu.2017.5145

Age at First Exposure to Repetitive Head Impacts Is Associated with Smaller Thalamic Volumes in Former Professional American Football Players

Vivian Schultz,[1,2] Robert A. Stern,[3,4] Yorghos Tripodis,[3,5] Julie Stamm,[1,3,6] Pawel Wrobel,[1,2] Christian Lepage,[1,7] Isabelle Weir,[5] Jeffrey P. Guenette,[1,8] Alicia Chua,[5] Michael L. Alosco,[3] Christine M. Baugh,[3,9] Nathan G. Fritts,[3] Brett M. Martin,[10] Christine E. Chaisson,[3,10] Michael J. Coleman,[1] Alexander P. Lin,[1,8,11] Ofer Pasternak,[1,8] Martha E. Shenton,[1,8,12,13,*] and Inga K. Koerte[1,2,13,*]

Abstract

Thalamic atrophy has been associated with exposure to repetitive head impacts (RHI) in professional fighters. The aim of this study is to investigate whether or not age at first exposure (AFE) to RHI is associated with thalamic volume in symptomatic former National Football League (NFL) players at risk for chronic traumatic encephalopathy (CTE). Eighty-six symptomatic former NFL players (mean age = 54.9 ± 7.9 years) were included. T1-weighted data were acquired on a 3T magnetic resonance imager, and thalamic volumes were derived using FreeSurfer. Mood and behavior, psychomotor speed, and visual and verbal memory were assessed. The association between thalamic volume and AFE to playing football and to number of years playing was calculated. Decreased thalamic volume was associated with more years of play (left: $p = 0.03$; right: $p = 0.03$). Younger AFE was associated with decreased right thalamic volume ($p = 0.014$). This association remained significant after adjusting for total years of play. Decreased left thalamic volume was associated with worse visual memory ($p = 0.014$), whereas increased right thalamic volume was associated with fewer mood and behavior symptoms ($p = 0.003$). In our sample of symptomatic former NFL players at risk for CTE, total years of play and AFE were associated with decreased thalamic volume. The effect of AFE on right thalamic volume was almost twice as strong as the effect of total years of play. Our findings confirm previous reports of an association between thalamic volume and exposure to RHI. They suggest further that younger AFE may result in smaller thalamic volume later in life.

Keywords: age at first exposure; chronic traumatic encephalopathy; repetitive head impacts; sports-related head injury; thalamus

Introduction

American football accounts for a high exposure to repetitive head impacts (RHI). During a single game, players may experience as many as 86 head impacts.[1] Chronic traumatic encephalopathy (CTE) is a neurodegenerative disease that has been associated with exposure to RHI.[2–5] Moreover, most of the pathologically confirmed cases of CTE have had a history of exposure to RHI, making RHI exposure a likely but not sufficient cause of CTE (for review, see [6]). Of further note, the clinical presentation of

[1]Psychiatry Neuroimaging Laboratory, Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.
[2]Department of Child and Adolescent Psychiatry, Psychosomatic, and Psychotherapy, Ludwig-Maximilian-University, Munich, Germany.
[3]BU Alzheimer's Disease and CTE Center, Boston University, Boston, Massachusetts.
[4]Departments of Neurology, Neurosurgery, and Anatomy & Neurobiology, Boston University School of Medicine, Boston, Massachusetts.
[5]Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts.
[6]Department of Kinesiology, University of Wisconsin, Madison, Madison, Wisconsin.
[7]Department of Psychology, University of Ottawa, Ottawa, Ontario, Canada.
[8]Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.
[9]Interfaculty Initiative in Health Policy, Harvard University, Boston, Massachusetts.
[10]Data Coordinating Center, Boston University School of Public Health, Boston, Massachusetts.
[11]Center for Clinical Spectroscopy, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.
[12]VA Boston Healthcare System, Brockton Division, Brockton, Massachusetts.
[13]Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts.
*Authors contributed equally.

CTE is not well understood.[7–9] Next-of-kin interviews and medical record reviews of deceased males with neuropathologically confirmed CTE suggest that CTE presents with impaired behavior (e.g., impulsivity, physical and verbal aggression), mood (e.g., depression, apathy, and related symptoms), and cognition (e.g., memory and executive dysfunction, and eventual dementia).[9] To date, CTE can only be diagnosed post-mortem. Thus, the development of biomarkers to support CTE diagnosis during life and the identification of contributing risk factors for the development of CTE are critical for early disease detection.[8]

In a study of neuropathologically confirmed cases of CTE in former football players, there was a significant association between total years playing and severity of post-mortem tau pathology.[10] In another study of living former National Football League (NFL) players, greater estimated cumulative RHI exposure (based, in part, on the total number of years played) was correlated with higher levels of later-life plasma tau.[11] A study of former high school and college football players demonstrated a dose-response relationship between estimated RHI exposure and risk of later-life impairments in mood, behavior, and cognition.[12]

There is evidence that not only exposure to RHI but also younger age at first exposure (AFE) may be an important risk factor for brain alterations in symptomatic former NFL players.[13,14] Stamm and associates[13,14] were the first to show the effect of AFE with playing football. Younger AFE (i.e., less than age 12) was associated with pronounced white matter alterations in the corpus callosum[14] as well as increased later-life cognitive impairment.[13] This vulnerability to RHI in children and adolescents may be linked closely to critical periods of brain development.[15,16] Despite the growing concern about long-term consequences of early exposure to RHI, however, 4.9 million youth athletes are participating currently in football in the United States.[17,18]

The thalamus is a central subcortical structure involved in the majority of cortex-controlled processes and pathways (for review see [19]). It is considered to be a key structure in cortical communication via nonreciprocal cortico-thalamo-cortical pathways, thereby providing an important hub for a dense flow of information between cortical areas.[20–22] Its prominent and distinctive position in the center of the brain along with strong connections to all parts of the cortex, as well as to numerous subcortical networks,[20] makes it a particularly important structure. Regional brain atrophy has been described in the thalamus and structures of the limbic system in CTE.[3,7,23] To date, however, only one study has investigated the association between thalamic volume and exposure to RHI.[24] In that study, Bernick and colleagues[24] examined 224 professional fighters using magnetic resonance imaging (MRI) and neuropsychological testing. Their results showed an association between decreased thalamic volume and two measures of exposure to RHI; total number of professional fights and years of professional fighting. In addition, the study reported an association between slower processing speed and smaller thalamic volumes.[24]

To what extent these findings are generalizable to athletes exposed to RHI while participating in contact sports other than professional fighting remains to be elucidated. Because the thalamus is a key structure in numerous functional networks,[25] whether and to what extent neurocognitive and behavioral function might be associated with thalamic volume warrants further investigation. To date, the relationship between RHI exposure, including the age when exposure begins, and thalamic volume in former professional football players has not been described. The age of first exposure is of particular public health interest given that far more youth athletes participate in football than those age 18 and older.[18] This study examined thalamic volume and its association with AFE, total years of play, as well as associations with neurobehavioral functioning, in symptomatic former professional football players.

Methods

This study was part of the Diagnosing and Evaluating Traumatic Encephalopathy using Clinical Tests (DETECT) project, funded by the National Institutes of Health (R01 NS 078337). The main goal of this project is to develop biomarkers for the *in vivo* diagnosis of CTE. Details of this project and its protocol are described elsewhere.[11,14,26,27] The study and related procedures were approved by the Boston University Medical Center Institutional Review Board and by the Partners Institutional Review Board. Written informed consent was obtained from all participants before enrollment.

Participants

Ninety-six former professional football players were included in the DETECT project. Study participants met the following inclusion criteria based on the DETECT study protocol: male, 40–69 years of age and a minimum of 12 years of organized football experience with at least two years of active participation in the NFL. Moreover, all participants had complaints of cognitive, mood, and behavioral symptoms for at least six months before participation, based on self-report. Potential participants with contraindications for MRI or lumbar puncture, history or diagnosis of any central nervous system disease, and English as a second language were excluded.

As part of DETECT, participants underwent a comprehensive assessment that included neurological as well as psychiatric examination, neuropsychological testing, standardized mood and behavior questionnaires, neuroimaging, a lumbar puncture for cerebrospinal fluid analysis, and genetic testing. The present study examines only imaging data and neuropsychological and mood and behavior measures.

Among the 96 enrolled subjects, 10 participants were excluded because of missing T1-weighted MRI data or poor data quality. Thus, a final sample size of 86 former NFL players was analyzed in this study (mean age: 54.9 ± 7.9 years*)*. For 75 former NFL players, both neurobehavioral and imaging data were available.

Exposure variables

The AFE and total years of play. The age participants started playing organized tackle football and the total number of years they played tackle football were reported. Both were treated as continuous variables.

MRI data acquisition

All neuroimaging data were acquired on a 3-Tesla MRI Scanner (Verio, Siemens Healthcare, Erlangen, Germany) with a 32-channel head array and the Syngo MR-B17 software suite. We acquired neuroimaging data on the 3-Tesla MRI Scanner with a 32-channel head array and the Syngo MR-B17 software suite. T1-weighted images were acquired with a three dimensional magnetization-prepared-rapid-gradient-echo sequence (MPRAGE): repetition time (TR) = 1800 msec, echo time (TE) = 3.36 msec, voxel size = $1 \times 1 \times 1$ mm^3, acquisition matrix = 256×256, flip angle = 7 degrees.

Image processing

Image data format was converted from DICOM to Nifti and visually inspected for quality. T1-weighted images were aligned and centered. Afterward, the T1-weighted images were automatically segmented using FreeSurfer 5.3 (http://surfer.nmr.mgh.harvard.edu/, Athinoula A. Martinos Center for Biomedical

Imaging, Charlestown, MA) resulting in a segmentation of the deep gray matter structures (including the thalamus) as well as a parcellation of the cortex.[28–30] The quality of the obtained FreeSurfer segmentation and parcellation was then visually assessed. An estimated total intracranial volume was automatically calculated for each participant using FreeSurfer. Finally, bilateral thalamic volume was calculated based on the FreeSurfer label maps of the thalamus in each hemisphere (Fig. 1).

Neurobehavioral outcomes: cognition, mood, and behavior

A set of standardized assessments was administered to examine cognitive function as well as mood and behavior.

Cognitive function. Participants completed the following measures of cognition: Trail Making Test A and B (TMT)[31]; Digit Symbol Coding from the Wechsler Adult Intelligence Scale - Revised (WAIS-R)[32]; Digit Span from the WAIS – R[32]; Wisconsin Card Sorting Test[33]; Animal Naming,[30] Controlled Oral Word Association Test (COWAT)[34]; Color-Word Interference Subtest from the Delis-Kaplan Executive Function System (DKEFS)[35]; Boston Qualitative Scoring System (BQSS) for the Rey-Osterrieth Complex Figure (ROCF)[36]; and the Story Learning, List Learning, Map Reading, and Naming Tests from the Neuropsychological Assessment Battery (NAB).[37] All tests were administered by a trained research assistant under the supervision of a licensed clinical neuropsychologist. All tests were double-scored by a second trained research assistant.

Mood and behavior. The following self-report measures were used to assess mood and behavior: Apathy Evaluation Scale (AES),[38] Barratt Impulsivity Scale (BIS-11),[39] Beck Depression Inventory II (BDI-II),[40] Beck Hopelessness Scale (BHS),[41] Behavior Rating Inventory of Executive Functioning - Adult Version (BRIEF-A),[42] Brown-Goodwin Lifetime History of Aggression (LHA),[43] Center for Epidemiologic Studies - Depression Scale (CES-D),[44] and the Buss-Durkee Inventory.[45] The Hamilton Depression Rating Scale (HDRS)[46] and the Modified Scale for Suicidal Ideation[47] were administered using a semi-structured interview by either a licensed psychiatrist or clinical psychologist.

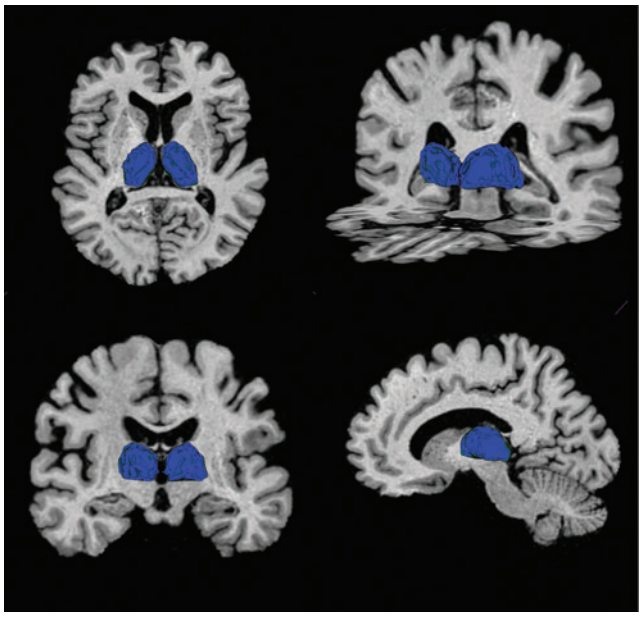


FIG. 1. A three-dimensional reconstruction of the thalamus as region of interest. The model was created from one randomly selected person using the model maker module of Slicer 4.5. Blue = thalamus. The model is shown from four different views on a para-mid axial, sagittal, and coronal slice and is superimposed on the individual T1-weighted image. The label map as basis for the three-dimensional model was not modified for the image creation.

Raw data from the neurobehavioral measures were converted to age-, gender-, and education-corrected standardized scores, when available. Next, on the basis of conceptual and empirical grounds, we generated four factor scores from the outcome measures using principal component analyses. Alosco and coworkers[48] provide a detailed account of our method for creating these factors. The factors and their constituent measures are as follows: Factor 1—Mood and Behavior: AES, BDI-II, BHS, BIS-11, BRIEF-A Behavioral Regulation Index, CES-D, HDRS, LHA; Factor 2—Attention and Psychomotor Speed: COWAT, DKEFS Color Word Interference (Inhibition/Switching score), TMT A and B, and Digit Symbol; Factor 3—Verbal Memory: NAB Story Learning (Phrase Unit Immediate and Delayed Recall scores) and NAB List Learning (Short and Long Delayed Recall scores); and Factor 4—Visual Memory: BQSS (Immediate Presence and Accuracy, and Delayed Presence and Accuracy scores).

Statistical analyses

We used Statistical Analysis System (SAS version 9.4; SAS Institute Inc., North Carolina) for all statistical analyses. Results of our inferential tests with a p value below 0.05 were reported as statistically significant. Both the mixed effects regression analyses exploring the relation between volume and exposure as well as the analyses investigating the association between volume and neurobehavioral factor scores were controlled for age, body mass index (BMI), and estimated total intracranial volume, adjusting for correlation of bilateral regions of interest. Moreover, the analyses exploring the relation between AFE and volume were adjusted for total years of play. Therefore, all presented results account for the positive relationship between AFE and total years of play.

To minimize variance, we employed a bootstrapping method.[49] Specifically, we resampled 500 replicates with replacement, each with a size equal to the original sample, and we reran the mixed effect-regression model across all replicates. The resulting confidence intervals were calculated using bias-adjustment correction.[50] We used the Mahalanobis distance to identify any extreme observations, but no outliers were found. Further, the variance was not found to increase with AFE. Similarly, we explored the associations between volume and the standardized neurobehavioral factor scores using mixed-effects regression models for all 75 participants with complete neurobehavioral and imaging data. We also computed partial correlations, adjusting for age, BMI, estimated total intracranial volume, and years of education.

Results

Table 1 summarizes the demographic and clinical characteristics of the sample. Mean age was 54.9 (standard deviation [SD] = 7.9) years. The AFE ranged from six to 17 years (mean: 11.8 years, SD = 2.6). The number of total years played ranged from 12 to 26 years (mean: 18.4 years, SD = 3.4).

Right and left thalamic volume was negatively correlated with total years of play (left, $p = 0.012$; right, $p = 0.03$; Table 2). The longer the time athletes actively participated in football, the smaller the thalamus. The right thalamus was on average 38.6 mm^3 and the left thalamus was on average 53.2 mm^3 smaller for every year of play. Right thalamic volume was associated with AFE after adjusting for total years of play ($p = 0.014$). Here, the younger the athlete's age

Table 1. Demographic and Clinical Characteristics

	Former professional football players (n = *86*)
	Mean (SD)
Age	54.86 (7.91)
AFE	11.77 (2.60)
Total years of play	18.41 (3.42)
Years of education	16.43 (0.96)
Body mass index	32.91 (4.96)
No. of concussions*	123.10 (580.00)
No. times lost consciousness	4.49 (16.50)

SD, standard deviation; AFE, age at first exposure.
*Based on self-report following being given a modern definition of concussion, as reported previously.[78,79]

when he started playing tackle football, the smaller his thalamic volume. In fact, AFE had a greater impact on thalamic volumes than did total years of play. That is, for every year a participant started playing earlier, the average decrease in thalamic volume was 64.9 mm^3 (**Fig. 2**). In contrast, there was no significant association between left thalamic volume and AFE ($p = 0.872$).

The relationships between the neurobehavioral factor scores and thalamic volume are presented in Table 3. A significant positive correlation was found between left thalamic volume and visual memory ($r = 0.28$, $p = 0.0143$). The smaller the left thalamus, the worse the individual's visual memory performance. Right thalamic volume was positively associated with mood and behavior—i.e., the worse the mood and behavioral symptoms, the larger the right thalamic volume ($r = 0.34$, $p = 0.0029$).

Discussion

The purpose of this study was to investigate the effect of exposure to repetitive concussive and subconcussive head impacts on thalamic volume, as well as the association between thalamic volume and neurobehavioral function in a group of symptomatic former professional football players, ages 40 to 69. Our results show that the longer an athlete participated in football and the younger he was when he began playing, the smaller the thalamic volume. Of note, the effect of the AFE to tackle football on right thalamic volume was almost twice as strong as the effect of total years of play.

The reasons for this loss in thalamic volume remain uncertain, although they might represent a manifestation of direct traumatic brain injury-associated pathology (including diffuse axonal injury-associated wallerian degeneration) or of neuroinflammatory and/or neurodegenerative pathologies.[51,52] Future studies are needed to investigate these potential mechanisms further.

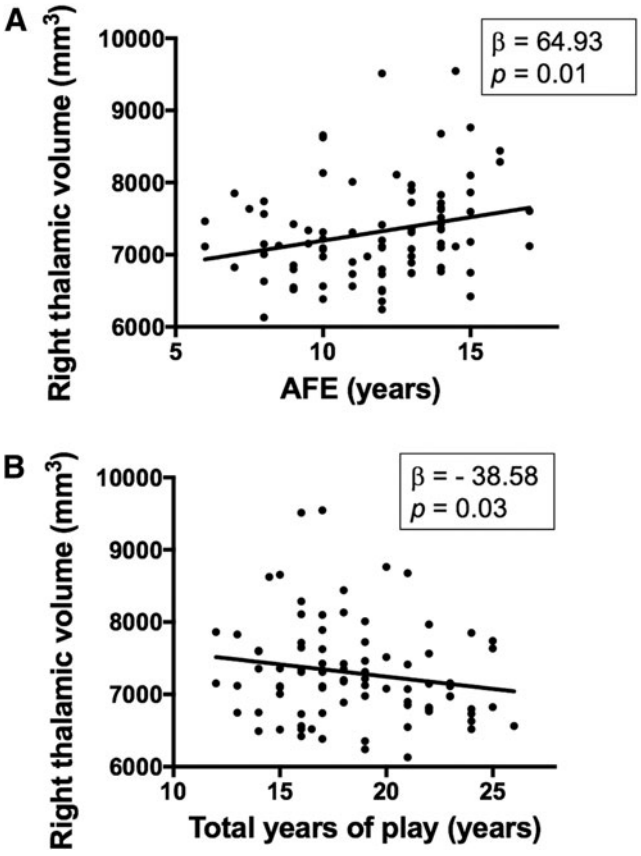


FIG. 2. Displays significant associations between exposure variables and right thalamic volume. (**A**) Association between age at first exposure (AFE) to football adjusted for total years of play and absolute right thalamic volume. (**B**) Relation between total years of play and absolute right thalamic volume.

Thalamic volume and exposure to RHI

Volume reduction in the thalamus has been reported in a variety of psychiatric and neurological disorders such as depression,[53] schizophrenia,[54] substance use disorder,[55] and in prodromal and later stages of Alzheimer disease.[56,57] Decreased thalamic volume has been reported in neuropathologically confirmed cases of CTE, which, in turn, was associated with a history of exposure to RHI.[3] Results from this study also confirm the literature on living athletes, where there is an association between decreased thalamic volume and concussion[58] and exposure to RHI.[24] More specifically, the study by Bernick and associates[24] of active professional fighters suggests that exposure to more fights is associated with decreased bilateral thalamic volumes. In fighters, thalamic volume decreased

Table 2. Thalamic Volumes and Exposure

Volume	*Exposure*	*Mean*	*Bias*	*Lower bound*	*Upper bound*	p *value*
Left thalamus	Total years of play	−53.2477	1.76261	−105.235	−7.80051	**0.012**
	AFE adjusted for total years of play	55.4455	0.80439	−62.3679	141.501	0.872
Right thalamus	Total years of play	−38.5831	0.44464	−86.486	−0.2263	**0.03**
	AFE adjusted for total years of play	64.9266	0.072667	2.2071	118.462	**0.014**

Bold indicates < 0.05.
AFE, age at first exposure.

Table 3. Thalamic Volumes and Neurobehavioral Factor Scores

	Left			*Right*		
	Partial correlation	t *value*	p *value*	*Partial correlation*	t *value*	p *value*
Factor 1: mood and behavior	0.14	1.23	0.2208	0.34	3.08	**0.0029**
Factor 2: attention/psychomotor speed	–0.03	–0.22	0.8245	–0.04	–0.34	0.7321
Factor 3: verbal memory	–0.19	–1.71	0.0912	–0.05	–0.4	0.6888
Factor 4: visual memory	0.28	2.51	**0.0143**	0.16	1.36	0.1777

Bold indicates < 0.05.

by 0.4% (right) and 0.3% (left) with each fight.[24] In a more recent study from the same group, Banks and colleagues[59] continued to find a statistically significant relationship between greater exposure and smaller thalamic volume. Moreover, apolipoprotein E genotype status did not impact this relationship.

Although linear head accelerations sustained during professional football are lower (60*g*)[60] than those sustained during boxing (71*g*)[61] and although exposure to RHI may vary between player positions in football,[1] our results are consistent with the aforementioned study in that exposure to RHI is associated with reduced thalamic volume. More specifically, total years of playing football were associated with a decrease in volume of 53.2 mm^3 in the left and 38.6 mm^3 in the right thalamus, per year of RHI exposure.

Thalamic volume and AFE

This is the first study showing that AFE to RHI in tackle football has an impact on subcortical gray matter volume. We found that younger AFE to playing tackle football was associated with a smaller right thalamic volume. In a previous study of the DETECT sample, younger AFE was associated with structural alterations in the corpus callosum.[14] In another study, younger AFE was associated with increased later-life cognitive impairment.[13] These results suggest that RHI may have a greater impact on the developing brain than on the mature brain.

Another study investigating AFE did not find significant results between RHI and brain structural changes. This study, however, examined brain structures that were not specific to developmental trajectories, and half of the sample did not play any youth football or played only one year of youth football.[62] Previous studies have also shown that age has a major influence on recovery from concussion, with adolescents experiencing more severe[63] and prolonged post-concussive deficits in visual and verbal memory[63–65] compared with adults.

This vulnerability to brain injury may be linked closely to a critical period of brain development and maturation during puberty. In fact, adolescence is considered a key period for brain maturation.[16] Important processes for proper brain development occurring during puberty are possibly selective and involve competitive elimination processes.[66,67] These processes include the downsizing of synapses, gray matter volume, glucose use, and neurotransmitter receptor densities.[68] The exact functional significance of these elimination patterns is unknown. They likely play a role in cortical plasticity necessary for accommodating to environmental needs that are crucial for brain maturation, however.[66,67] Moreover, these developmental processes do not occur uniformly.[68,69] In fact, maturation of the cerebral cortex seems to follow a hetero-chronic pattern that shows regional variation.[16,68–70] This means that age of onset for these processes likely vary between brain structures, resulting in a period of several years during adolescence in which the brain may be particularly vulnerable to external insults.[71]

Earlier age for starting to play football could be linked possibly to more years of play and thus have driven the results. It is therefore noteworthy that the association between thalamic volume and AFE remained significant when adjusting for total years of play. In fact, the effect of AFE on right thalamic volume was almost double the effect of total years of play with a decrease of 64.9 mm^3 for every year athletes had started playing earlier and a decrease of 38.6 mm^3 for every year athletes had played longer. The significant association between AFE and thalamic volume emphasizes that every year of difference in AFE accounts for reduced thalamic volume later in life.

Previous studies[13,14] report AFE below age 12 to be significantly associated with increased impairment and structural brain alterations. The linear relationship found here, however, does not contradict a threshold at a certain age as found in previous studies.[13,14] This difference with previous results may be explained by the high between-subject variability in thalamic volume, which decreases power to detect significant thresholds at a certain age. It could also be because of differing neurodevelopmental trajectories in the investigated brain regions. For example, Brown and coworkers[71] showed that the thalamus increases in volume until the age of 17.8 years with a plateau reached at around age 20, whereas another subcortical gray matter structure, the hippocampus, already peaked by 14.2 years and decreased in volume thereafter.[71] These examples demonstrate the importance of taking into account the developmental trajectory of the structure and/or function being studied with regard to AFE.

Importantly, only the right thalamus seemed to be affected by AFE. A possible explanation here is the difference in development between the two hemispheres. The left hemisphere has a peak developmental time approximately between age three and six, whereas the right hemisphere has a growth spurt between age eight and 10.[72] The latter is much more likely to reflect interference in neurodevelopmental processes as children begin playing football closer in age to this growth spurt than to that of the left hemisphere (mean AFE was 11.8 years in our study cohort).

Thalamic volume and neurobehavioral functioning

This study also revealed a significant association between decreased left thalamic volume and worse visual memory performance. This observation is consistent with a previous report by Konstantinou and colleagues[73] who examined 17 male patients after an average of 8.36 years of moderate to severe traumatic brain injury. They used neuroimaging as well as a neuropsychological testing battery to investigate the relation between structural brain alterations and neurocognitive outcome measures. Smaller thalamic volumes correlated with lower scores in visual memory, as assessed

with the ROCF (the same test used in our study), as well as with all other neurocognitive measures. Thalamic pathology has also been implicated in early memory loss in neurodegenerative disorders, including Alzheimer's disease.[74] It is not clear, however, why there was only a relationship between thalamic volume and visual but not verbal memory.

Mood and behavior were positively correlated with right thalamic volume, meaning that the larger the thalamic volume, the worse the reported mood and behavioral symptoms. It is noteworthy that larger right thalamic volume has been found in medication-naïve depressive patients[75] and that a decrease in thalamic volumes has been associated with successful antidepressant treatment.[76] Thus, our finding of an association between right thalamic volume and worse mood and behavioral symptoms may reflect the previously described relationship between thalamic volume and depressive symptoms. Future studies are nonetheless needed to further investigate this association.

Limitations

The cross-sectional study design precluded a more definitive link between RHI and thalamic volume as well as the establishment of causality between AFE and thalamic atrophy. Future studies should include a longitudinal design to understand better RHI-related brain changes as well as neurobehavioral changes. Moreover, future studies should investigate anatomical subsegments of the thalamus separately. Another limitation of this study is the lack of a comparison group of asymptomatic former professional football players. Based on the study design that only symptomatic former professional football players were included in DETECT, the results presented here are not generalizable to asymptomatic former professional football players or football players who did not compete at a professional level.

A further limitation of our study is that our findings are not generalizable to football players who did not compete at a professional level. The results may also not generalize to other groups frequently exposed to RHI. Future studies should include asymptomatic former professional football players as well as male and female athletes participating in different levels of intensity and in different contact sports. Finally, without post-mortem analysis, it is not possible to determine whether or not our findings are specifically related to the tauopathy of CTE or to other long-term consequences of RHI—e.g., among the reasons for a smaller thalamic volume may be exposure to RHI (i.e., microgliosis[77]).

Conclusion

Younger AFE and duration of exposure to RHI are associated with smaller thalamic volume in symptomatic former professional football players, a population at increased risk for the development of CTE. Smaller thalamic volume is also associated with worse performance in visual memory tasks. Longitudinal studies are needed to confirm these findings as well as to examine the association between neurodevelopmental processes and exposure to RHI.

Acknowledgments

The authors extend their appreciation to the study participants who make this work possible. This study was supported by the NIH R01 NS 078337 (RAS, YT, CMB, NGF, BMM, CC); F31 NS 081957 (JMS); 1F32NS096803-01 (MLA); P30 AG13846 (RAS, YT, MLA, CMB, BMM, CC); UL1-TR000157 (RAS); P41 EB015902 (OP); T32GM074905 (IW), and participant travel was partially funded by JetBlue Airlines, the National Football League, and the NFL Players Association. The sponsors had no role in the design or conduct of the study, the collection, analysis, or interpretation of the data, in the preparation, review, or approval of the manuscript, or in the decision to submit the study for publication. This study was also partly supported by the Else Kröner-Fresenius Foundation, Germany (IK), by a VA Merit Award (MES, MC), by the Canadian Institutes of Health Research Frederick Banting and Charles Best Doctoral Award (CL), and by the German Academic Exchange Service PROMOS award (VS). Finally, this work was completed in partial fulfillment of VS's dissertation.

Author Disclosure Statement

RAS is a paid consultant to Avanir Pharmaceuticals, Inc. (Aliso Viejo, CA), Biogen (Cambridge, MA), and Eli Lilly (Indianapolis, IN). He receives royalties for published neuropsychological tests (including the NAB and BQSS) from Psychological Assessment Resources, Inc. (Lutz, FL). CMB receives research funding through the Harvard Football Players Health Study, which is funded by the NFL Players' Association. For all other authors, no competing financial interests exist.

References

1. Crisco, J.J., Fiore, R., Beckwith, J.G., Chu, J.J., Brolinson, P.G., Duma, S., McAllister, T.W., Duhaime, A.C., and Greenwald, R.M. (2010). Frequency and location of head impact exposures in individual collegiate football players. J. Athl. Train. 45, 549–559.
2. Lenihan, M.W., and Jordan, B.D. (2015). The clinical presentation of chronic traumatic encephalopathy. Curr. Neurol. Neurosci. Rep. 15, 23.
3. McKee, A.C., Cantu, R.C., Nowinski, C.J., Hedley-Whyte, E.T., Gavett, B.E., Budson, A.E., Santini, V.E., Lee, H.S., Kubilus, C.A., and Stern, R.A. (2009). Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury. J. Neuropathol. Exp. Neurol. 68, 709–735.
4. Mez, J., Stern, R.A., and McKee, A.C. (2013). Chronic traumatic encephalopathy: where are we and where are we going? Curr. Neurol. Neurosci. Rep. 13, 407.
5. Ling, H., Morris, H.R., Neal, J.W., Lees, A.J., Hardy, J., Holton, J.L., Revesz, T., and Williams, D.D. (2017). Mixed pathologies including chronic traumatic encephalopathy account for dementia in retired association football (soccer) players. Acta Neuropathol. 133, 337–352.
6. Baugh, C.M., Stamm, J.M., Riley, D.O., Gavett, B.E., Shenton, M.E., Lin, A., Nowinski, C.J., Cantu, R.C., McKee, A.C., and Stern, R.A. (2012). Chronic traumatic encephalopathy: neurodegeneration following repetitive concussive and subconcussive brain trauma. Brain Imaging Behav. 6, 244–254.
7. McKee, A.C., Stern, R.A., Nowinski, C.J., Stein, T.D., Alvarez, V.E., Daneshvar, D.H., Lee, H.S., Wojtowicz, S.M., Hall, G., Baugh, C.M., Riley, D.O., Kubilus, C.A., Cormier, K.A., Jacobs, M.A., Martin, B.R., Abraham, C.R., Ikezu, T., Reichard, R.R., Wolozin, B.L., Budson, A.E., Goldstein, L.E., Kowall, N.W., and Cantu, R.C. (2013). The spectrum of disease in chronic traumatic encephalopathy. Brain 136, 43–64.
8. Montenigro, P.H., Baugh, C.M., Daneshvar, D.H., Mez, J., Budson, A.E., Au, R., Katz, D.I., Cantu, R.C., and Stern, R.A. (2014). Clinical subtypes of chronic traumatic encephalopathy: literature review and proposed research diagnostic criteria for traumatic encephalopathy syndrome. Alzheimers Res. Ther. 6, 68.
9. Stern, R.A., Daneshvar, D.H., Baugh, C.M., Seichepine, D.R., Montenigro, P.H., Riley, D.O., Fritts, N.G., Stamm, J.M., Robbins, C.A., McHale, L., Simkin, I., Stein, T.D., Alvarez, V.E., Goldstein, L.E., Budson, A.E., Kowall, N.W., Nowinski, C.J., Cantu, R.C., and McKee, A.C. (2013). Clinical presentation of chronic traumatic encephalopathy. Neurology 81, 1122–1129.
10. Cherry, J.D., Tripodis, Y., Alvarez, V.E., Huber, B., Kiernan, P.T., Daneshvar, D.H., Mez, J., Montenigro, P.H., Solomon, T.M., Alosco, M.L., Stern, R.A., McKee, A.C., and Stein, T.D. (2016). Microglial neuroinflammation contributes to tau accumulation in chronic traumatic encephalopathy. Acta Neuropathol. Commun. 4, 112.

11. Alosco, M.L., Tripodis, Y., Jarnagin, J., Baugh, C.M., Martin, B., Chaisson, C.E., Estochen, N., Song, L., Cantu, R.C., Jeromin, A., and Stern, R.A. (2017). Repetitive head impact exposure and later-life plasma total tau in former National Football League players. Alzheimers Dement. (Amst) 7, 33–40.
12. Montenigro, P.H., Alosco, M.L., Martin, B.M., Daneshvar, D.H., Mez, J., Chaisson, C.E., Nowinski, C.J., Au, R., McKee, A.C., Cantu, R.C., McClean, M.D., Stern, R.A., and Tripodis, Y. (2017). Cumulative head impact exposure predicts later-life depression, apathy, executive dysfunction, and cognitive impairment in former high school and college football players. J. Neurotrauma 34, 328–340.
13. Stamm, J.M., Bourlas, A.P., Baugh, C.M., Fritts, N.G., Daneshvar, D.H., Martin, B.M., McClean, M.D., Tripodis, Y., and Stern, R.A. (2015). Age of first exposure to football and later-life cognitive impairment in former NFL players. Neurology 84, 1114–1120.
14. Stamm, J.M., Koerte, I.K., Muehlmann, M., Pasternak, O., Bourlas, A.P., Baugh, C.M., Giwerc, M.Y., Zhu, A., Coleman, M.J., Bouix, S., Fritts, N.G., Martin, B.M., Chaisson, C., McClean, M.D., Lin, A.P., Cantu, R.C., Tripodis, Y., Stern, R.A., and Shenton, M.E. (2015). Age at first exposure to football is associated with altered corpus callosum white matter microstructure in former professional football players. J. Neurotrauma 32, 1768–1776.
15. Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., and Rapoport, J.L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. Nat. Neurosci. 2, 861–863.
16. Lebel, C., Walker, L., Leemans, A., Phillips, L., and Beaulieu, C. (2008). Microstructural maturation of the human brain from childhood to adulthood. Neuroimage 40, 1044–1055.
17. Powell, J.W., and Barber-Foss, K.D. (1999). Traumatic brain injury in high school athletes. JAMA 282, 958–963.
18. Daniel, R.W., Rowson, S., and Duma, S.M. (2012). Head impact exposure in youth football. Ann. Biomed. Eng. 40, 976–981.
19. Bell, P.T., and Shine, J.M. (2016). Subcortical contributions to large-scale network communication. Neurosci. Biobehav. Rev. 71, 313–322.
20. Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., Mortrud, M.T., Ouellette, B., Nguyen, T.N., Sorensen, S.A., Slaughterbeck, C.R., Wakeman, W., Li, Y., Feng, D., Ho, A., Nicholas, E., Hirokawa, K.E., Bohn, P., Joines, K.M., Peng, H., Hawrylycz, M.J., Phillips, J.W., Hohmann, J.G., Wohnoutka, P., Gerfen, C.R., Koch, C., Bernard, A., Dang, C., Jones, A.R., and Zeng, H. (2014). A mesoscale connectome of the mouse brain. Nature 508, 207–214.
21. Saalmann, Y.B., Pinsk, M.A., Wang, L., Li, X., and Kastner, S. (2012). The pulvinar regulates information transmission between cortical areas based on attention demands. Science 337, 753–756.
22. Sherman, S.M. (2007). The thalamus is more than just a relay. Curr. Opin. Neurobiol. 17, 417–422.
23. Montenigro, P.H., Corp, D.T., Stein, T.D., Cantu, R.C., and Stern, R.A. (2015). Chronic traumatic encephalopathy: historical origins and current perspective. Annu. Rev. Clin. Psychol. 11, 309–330.
24. Bernick, C., Banks, S.J., Shin, W., Obuchowski, N., Butler, S., Noback, M., Phillips, M., Lowe, M., Jones, S., and Modic, M. (2015). Repeated head trauma is associated with smaller thalamic volumes and slower processing speed: the Professional Fighters' Brain Health Study. Br. J. Sports Med. 49, 1007–1011.
25. Herrero, M.T., Barcia, C., and Navarro, J.M. (2002). Functional anatomy of thalamus and basal ganglia. Childs Nerv. Syst. 18, 386–404.
26. Koerte, I.K., Hufschmidt, J., Muehlmann, M., Tripodis, Y., Stamm, J.M., Pasternak, O., Giwerc, M.Y., Coleman, M.J., Baugh, C.M., Fritts, N.G., Heinen, F., Lin, A., Stern, R.A., and Shenton, M.E. (2016). Cavum septi pellucidi in symptomatic former professional football players. J. Neurotrauma 33, 346–353.
27. Stern, R.A., Tripodis, Y., Baugh, C.M., Fritts, N.G., Martin, B.M., Chaisson, C., Cantu, R.C., Joyce, J.A., Shah, S., Ikezu, T., Zhang, J., Gercel-Taylor, C., and Taylor, D.D. (2016). Preliminary study of plasma exosomal tau as a potential biomarker for chronic traumatic encephalopathy. J. Alzheimers Dis. 51, 1099–1109.
28. Desikan, R.S., Segonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., and Killiany, R.J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31, 968–980.
29. Fischl, B., Salat, D.H., van der Kouwe, A.J., Makris, N., Segonne, F., Quinn, B.T., and Dale, A.M. (2004). Sequence-independent segmentation of magnetic resonance images. Neuroimage 23, Suppl 1, S69–S84.
30. Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., and Dale, A.M. (2004). Automatically parcellating the human cerebral cortex. Cereb. Cortex 14, 11–22.
31. Reitan, R.M. (1992). *Trail Making Test: Manual for Administration and Scoring.* Reitan Neuropsychology Laboratory: Tucson, AZ.
32. Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale-III.* Psychological Corporation: San Antonio, TX.
33. Heaton, R.K., Chelune, G.J., Talley, J., Kay, G.G., and Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and Expanded.* Psychological Assessment Resources Inc: Odessa, FL.
34. Lezak, M.D., Howieson, D.B., Bigler, E.D., and Tranel, D. (2012). *Neuropsychological Assessment,* 5th ed. Oxford University Press: Oxford, UK.
35. Delis, D.C., Kaplan, E., and Kramer, J.H. (2001). *Delis-Kaplan Executive Function System (D-KEFS).* Psychological Corporation: San Antonio, TX.
36. Stern, R.A., Javorsky, D.J., Singer, E.A., Singer, N.G., Duke, L.M., Somerville, J.A., Thompson, J.A., and Kaplan, E. (1999). *Boston Qualitative Scoring System (BQSS) for the Rey-Osterrieth Complex Figure.* Psychological Assessment Resources (PAR): Odessa, FL.
37. Stern, R., and White, T. (2003). *The Neuropsychological Assessment Battery (NAB).* Psychological Assessment Resources: Lutz, Florida.
38. Marin, R.S., Biedrzycki, R.C., and Firinciogullari, S. (1991). Reliability and validity of the Apathy Evaluation Scale, Psychiatry Res. 38, 143–162.
39. Barratt, E.S. (1965). Factor analysis of some psychometric measures of impulsiveness and anxiety. Psychol. Rep. 16, 547–554.
40. Beck, A., Steer, R., and Brown, G. (1996). *Manual for the BDI-II.* Psychological Corporation: San Antonio, TX.
41. Beck, A., and Steer, R. (1988). *Manual for the Beck Hopelessness Scale.* Psychological Corporation: San Antonio, TX.
42. Roth, R.M., Isquith, P.K., and Gioia, G.A. (2005). *BRIEF-A: Behavior Rating Inventory of Executive Function*—adult version: professional manual. Psychological Assessment Resources: Lutz, Florida.
43. Brown, G.L., Goodwin, F.K., Ballenger, J.C., Goyer, P.F., and Major, L.F. (1979). Aggression in humans correlates with cerebrospinal fluid amine metabolites. Psychiatry Res. 1, 131–139.
44. Radloff, L.S. (1977). The CES-D Scale: a self-report depression scale for research in the general population. Applied Psychological Measurement. 1, 385–401.
45. Buss, A.H., and Durkee, A. (1957). An inventory for assessing different kinds of hostility. J. Consult. Psychol. 21, 343–349.
46. Hamilton, M. (1960). A rating scale for depression. J. Neurol. Neurosurg. Psychiatry 23, 56–62.
47. Miller, I.W., Norman, W.H., Bishop, S.B., and Dow, M.G. (1986). The Modified Scale for Suicidal Ideation: reliability and validity. J. Consult. Clin. Psychol. 54, 724–725.
48. Alosco, M.L., Jarnagin, J., Tripodis, Y., Platt, M., Martin, B., Chaisson, C.E., Baugh, C.M., Fritts, N.G., Cantu, R.C., and Stern, R.A. (2017). Olfactory function and associated clinical correlates in former National Football League players. J. Neurotrauma 34, 772–780.
49. Efron, B., and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap.* Chapman & Hall/CRC Press: Boca Raton, FL.
50. Efron, B. (1987). Better bootstrap confidence intervals. J. Am. Stat. Assoc. 82, 171–185.
51. Johnson, V.E., Stewart, W., and Smith, D.H. (2013). Axonal pathology in traumatic brain injury. Exp. Neurol. 246, 35–43.
52. Koerte, I.K., Lin, A.P., Muehlmann, M., Merugumala, S., Liao, H., Starr, T., Kaufmann, D., Mayinger, M., Steffinger, D., Fisch, B., Karch, S., Heinen, F., Ertl-Wagner, B., Reiser, M., Stern, R.A., Zafonte, R., and Shenton, M.E. (2015). Altered neurochemistry in former professional soccer players without a history of concussion. J. Neurotrauma 32, 1287–1293.
53. Webb, C.A., Weber, M., Mundy, E.A., and Killgore, W.D. (2014). Reduced gray matter volume in the anterior cingulate, orbitofrontal cortex and thalamus as a function of mild depressive symptoms: a voxel-based morphometric analysis. Psychol. Med. 44, 2833–2843.
54. Okada, N., Fukunaga, M., Yamashita, F., Koshiyama, D., Yamamori, H., Ohi, K., Yasuda, Y., Fujimoto, M., Watanabe, Y., Yahata, N.,

Nemoto, K., Hibar, D.P., van Erp, T.G., Fujino, H., Isobe, M., Isomura, S., Natsubori, T., Narita, H., Hashimoto, N., Miyata, J., Koike, S., Takahashi, T., Yamasue, H., Matsuo, K., Onitsuka, T., Iidaka, T., Kawasaki, Y., Yoshimura, R., Watanabe, Y., Suzuki, M., Turner, J.A., Takeda, M., Thompson, P.M., Ozaki, N., Kasai, K., and Hashimoto, R. (2016). Abnormal asymmetries in subcortical brain volume in schizophrenia. Mol. Psychiatry 21, 1460–1466.
55. Grodin, E.N., and Momenan, R. (2016). Decreased subcortical volumes in alcohol dependent individuals: effect of polysubstance use disorder. Addict. Biol. 22, 1426–1437.
56. Eustache, P., Nemmi, F., Saint-Aubert, L., Pariente, J., and Péran, P. (2016). Multimodal magnetic resonance imaging in Alzheimer's disease patients at prodromal stage. Addict. Biol. 50, 1035–1050.
57. Zarei, M., Patenaude, B., Damoiseaux, J., Morgese, C., Smith, S., Matthews, P.M., Barkhof, F., Rombouts, S.A., Sanz-Arigita, E., and Jenkinson, M. (2010). Combining shape and connectivity analysis: an MRI study of thalamic degeneration in Alzheimer's disease. Neuroimage 49, 1–8.
58. Zagorchev, L., Meyer, C., Stehle, T., Wenzel, F., Young, S., Peters, J., Weese, J., Paulsen, K., Garlinghouse, M., Ford, J., Roth, R., Flashman, L., and McAllister, T. (2016). Differences in regional brain volumes two months and one year after mild traumatic brain injury. J. Neurotrauma 33, 29–34.
59. Banks, S.J., Miller, J.B., Rissman, R.A., and Bernick, C.B. (2017). Lack of influence of apolipoprotein e status on cognition or brain structure in professional fighters. J. Neurotrauma 34, 380–384.
60. Pellman, E.J., Viano, D.C., Tucker, A.M., Casson, I.R., and Waeckerle, J.F. (2003). Concussion in professional football: reconstruction of game impacts and injuries. Neurosurgery 53, 799–812.
61. Viano, D.C., Casson, I.R., Pellman, E.J., Bir, C.A., Zhang, L., Sherman, D.C., and Boitano, M.A. (2005). Concussion in professional football: comparison with boxing head impacts—part 10. Neurosurgery 57, 1154–1172.
62. Solomon, G.S., Kuhn, A.W., Zuckerman, S.L., Casson, I.R., Viano, D.C., Lovell, M.R., and Sills, A.K. (2016). Participation in pre-high school football and neurological, neuroradiological, and neuropsychological findings in later life: a study of 45 retired National Football League players. Am. J. Sports Med. 44, 1106–1115.
63. Covassin, T., Elbin, R.J., Harris, W., Parker, T., and Kontos, A. (2012). The role of age and sex in symptoms, neurocognitive performance, and postural stability in athletes after concussion. Am. J. Sports Med. 40, 1303–1312.
64. Field, M., Collins, M.W., Lovell, M.R., and Maroon, J. (2003). Does age play a role in recovery from sports-related concussion? A comparison of high school and collegiate athletes. J. Pediatr. 142, 546–553.
65. Zuckerman, S.L., Lee, Y.M., Odom, M.J., Solomon, G.S., Forbes, J.A., and Sills, A.K. (2012). Recovery from sports-related concussion: days to return to neurocognitive baseline in adolescents versus young adults. Surg. Neurol. Int. 3, 130.
66. Spear, L.P. (2000). The adolescent brain and age-related behavioral manifestations. Neurosci. Biobehav. Rev. 24, 417–463.
67. Bourgeois, J.P., Jastreboff, P.J., and Rakic, P. (1989). Synaptogenesis in visual cortex of normal and preterm monkeys: evidence for intrinsic regulation of synaptic overproduction. Proc. Natl. Acad. Sci. U. S. A. 86, 4297–4301.
68. Giedd, J.N. (2008). The teen brain: insights from neuroimaging. J. Adolesc. Health 42, 335–343.
69. Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent, T.F., 3rd, Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L., and Thompson, P.M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. Proc. Natl. Acad. Sci. U. S. A. 101, 8174–8179.
70. Whitford, T.J., Rennie, C.J., Grieve, S.M., Clark, C.R., Gordon, E., and Williams, L.M. (2007). Brain maturation in adolescence: concurrent changes in neuroanatomy and neurophysiology. Hum. Brain Mapp. 28, 228–237.
71. Brown, T.T., Kuperman, J.M., Chung, Y., Erhart, M., McCabe, C., Hagler, D.J., Jr., Venkatraman, V.K., Akshoomoff, N., Amaral, D.G., Bloss, C.S., Casey, B.J., Chang, L., Ernst, T.M., Frazier, J.A., Gruen, J.R., Kaufmann, W.E., Kenet, T., Kennedy, D.N., Murray, S.S., Sowell, E.R., Jernigan, T.L., and Dale, A.M. (2012). Neuroanatomical assessment of biological maturity. Curr. Biol. 22, 1693–1698.
72. Thatcher, R.W., Walker, R.A., and Giudice, S. (1987). Human cerebral hemispheres develop at different rates and ages. Science 236, 1110–1113.
73. Konstantinou, N., Pettemeridou, E., Seimenis, I., Eracleous, E., Papacostas, S.S., Papanicolaou, A.C., and Constantinidou, F. (2016). Assessing the relationship between neurocognitive performance and brain volume in chronic moderate-severe traumatic brain injury. Front. Neurol. 7, 29.
74. Aggleton, J.P., Pralus, A., Nelson, A.J., and Hornberger, M. (2016). Thalamic pathology and memory loss in early Alzheimer's disease: moving the focus from the medial temporal lobe to Papez circuit. Brain 139, 1877–1890.
75. Zhao, Y.J., Du, M.Y., Huang, X.Q., Lui, S., Chen, Z.Q., Liu, J., Luo, Y., Wang, X.L., Kemp, G.J., and Gong, Q.Y. (2014). Brain grey matter abnormalities in medication-free patients with major depressive disorder: a meta-analysis. Psychol. Med. 44, 2927–2937.
76. Young, K.A., Bonkale, W.L., Holcomb, L.A., Hicks, P.B., and German, D.C. (2008). Major depression, 5HTTLPR genotype, suicide and antidepressant influences on thalamic volume. Br. J. Psychiatry 192, 285–289.
77. Robinson, S., Berglass, J.B., Denson, J.L., Berkner, J., Anstine, C.V., Winer, J.L., Maxwell, J.R., Qiu, J., Yang, Y., Sillerud, L.O., Meehan, W.P., 3rd, Mannix, R., and Jantzie, L.L. (2017). Microstructural and microglial changes after repetitive mild traumatic brain injury in mice. J. Neurosci. Res. 95, 1025–1035.
78. Robbins, C.A., Daneshvar, D.H., Picano, J.D., Gavett, B.E., Baugh, C.M., Riley, D.O., Nowinski, C.J., McKee, A.C., Cantu, R.C., and Stern, R.A. (2014). Self-reported concussion history: impact of providing a definition of concussion. Open Access J. Sports Med. 5, 99–103.
79. Alosco, M.L., Jarnagin, J., Tripodis, Y., Martin, B., Chaisson, C., Baugh, C.M., Torres, A., Nowinski, C.J., Cantu, R.C., and Stern, R.A. (2017). Utility of providing a concussion definition in the assessment of concussion history in former NFL players. Brain Inj. 31, 1116–1123.

Address correspondence to:
Inga K. Koerte, MD
Psychiatry Neuroimaging Laboratory
1249 Boylston Street
Boston, MA 02215

E-mail: ikoerte@bwh.harvard.edu

REFERENCES

1. Lenihan MW, Jordan BD. The clinical presentation of chronic traumatic encephalopathy. *Curr Neurol Neurosci Rep.* 2015;15(5):23.
2. McKee AC, Cantu RC, Nowinski CJ, et al. Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury. *J Neuropathol Exp Neurol.* 2009;68(7):709-735.
3. Mez J, Stern RA, McKee AC. Chronic traumatic encephalopathy: where are we and where are we going? *Curr Neurol Neurosci Rep.* 2013;13(12):407.
4. Martland HS. Punch Drunk. *Journal of the American Medical Association.* 1928;91(15).
5. Millspaugh J. Dementia pugilistica. *US Naval Med Bulletin.* 1937;35:297–261.
6. Critchley M. Punch-drunk syndromes: the chronic traumatic encephalopathy of boxers. *Hommage à Clovis Vincent*. Paris: Maloine1949.
7. Bailes JE, Petraglia AL, Omalu BI, Nauman E, Talavage T. Role of subconcussion in repetitive mild traumatic brain injury. *J Neurosurg.* 2013;119(5):1235-1245.
8. Baugh CM, Stamm JM, Riley DO, et al. Chronic traumatic encephalopathy: neurodegeneration following repetitive concussive and subconcussive brain trauma. *Brain Imaging Behav.* 2012;6(2):244-254.
9. Montenigro PH, Baugh CM, Daneshvar DH, et al. Clinical subtypes of chronic traumatic encephalopathy: literature review and proposed research diagnostic criteria for traumatic encephalopathy syndrome. *Alzheimers Res Ther.* 2014;6(5):68.
10. Mez J, Daneshvar DH, Kiernan PT, et al. Clinicopathological Evaluation of Chronic Traumatic Encephalopathy in Players of American Football. *JAMA.* 2017;318(4):360-370.
11. Gavett BE, Stern RA, McKee AC. Chronic traumatic encephalopathy: a potential late effect of sport-related concussive and subconcussive head trauma. *Clin Sports Med.* 2011;30(1):179-188, xi.
12. Geddes JF, Vowles GH, Nicoll JA, Revesz T. Neuronal cytoskeletal changes are an early consequence of repetitive head injury. *Acta Neuropathol.* 1999;98(2):171-178.
13. Hof PR, Knabe R, Bovier P, Bouras C. Neuropathological observations in a case of autism presenting with self-injury behavior. *Acta Neuropathol.* 1991;82(4):321-326.
14. Omalu BI, DeKosky ST, Hamilton RL, et al. Chronic traumatic encephalopathy in a national football league player: part II. *Neurosurgery.* 2006;59(5):1086-1092; discussion 1092-1083.
15. Roberts GW, Whitwell HL, Acland PR, Bruton CJ. Dementia in a punch-drunk wife. *Lancet.* 1990;335(8694):918-919.
16. Stern RA, Daneshvar DH, Baugh CM, et al. Clinical presentation of chronic traumatic encephalopathy. *Neurology.* 2013;81(13):1122-1129.
17. Goldstein LE, Fisher AM, Tagge CA, et al. Chronic traumatic encephalopathy in blast-exposed military veterans and a blast neurotrauma mouse model. *Sci Transl Med.* 2012;4(134):134ra160.
18. Gooch CL, Pracht E, Borenstein AR. The burden of neurological disease in the United States: A summary report and call to action. *Ann Neurol.* 2017;81(4):479-484.
19. McKee AC, Stern RA, Nowinski CJ, et al. The spectrum of disease in chronic traumatic encephalopathy. *Brain.* 2013;136(Pt 1):43-64.
20. McKee AC, Cairns NJ, Dickson DW, et al. The first NINDS/NIBIB consensus meeting to define neuropathological criteria for the diagnosis of chronic traumatic encephalopathy. *Acta Neuropathol.* 2016;131(1):75-86.
21. Huber BR, Alosco ML, Stein TD, McKee AC. Potential Long-Term Consequences of Concussive and Subconcussive Injury. *Phys Med Rehabil Clin N Am.* 2016;27(2):503-511.
22. McKee AC, Daneshvar DH. *The neuropathology of traumatic brain injury.* Vol 127. 2015/02/24 ed2015.

23. Guo T, Noble W, Hanger DP. Roles of tau protein in health and disease. *Acta Neuropathol.* 2017;133(5):665-704.
24. Lu KP, Kondo A, Albayram O, Herbert MK, Liu H, Zhou XZ. Potential of the Antibody Against cis-Phosphorylated Tau in the Early Diagnosis, Treatment, and Prevention of Alzheimer Disease and Brain Injury. *JAMA Neurol.* 2016;73(11):1356-1362.
25. Kondo A, Shahpasand K, Mannix R, et al. Antibody against early driver of neurodegeneration cis P-tau blocks brain injury and tauopathy. *Nature.* 2015;523(7561):431-436.
26. Meehan W, 3rd, Mannix R, Zafonte R, Pascual-Leone A. Chronic traumatic encephalopathy and athletes. *Neurology.* 2015;85(17):1504-1511.
27. Montenigro PH, Corp DT, Stein TD, Cantu RC, Stern RA. Chronic traumatic encephalopathy: historical origins and current perspective. *Annu Rev Clin Psychol.* 2015;11:309-330.
28. Perrine K, Helcer J, Tsiouris AJ, Pisapia DJ, Stieg P. The Current Status of Research on Chronic Traumatic Encephalopathy. *World Neurosurg.* 2017;102:533-544.
29. Jordan BD. The clinical spectrum of sport-related traumatic brain injury. *Nat Rev Neurol.* 2013;9(4):222-230.
30. Reams N, Eckner JT, Almeida AA, et al. A Clinical Approach to the Diagnosis of Traumatic Encephalopathy Syndrome: A Review. *JAMA Neurol.* 2016;73(6):743-749.
31. Victoroff J. Traumatic encephalopathy: review and provisional research diagnostic criteria. *NeuroRehabilitation.* 2013;32(2):211-224.
32. Koerte IK, Hufschmidt J, Muehlmann M, et al. Cavum Septi Pellucidi in Symptomatic Former Professional Football Players. *J Neurotrauma.* 2016;33(4):346-353.
33. Ng TS, Lin AP, Koerte IK, et al. Neuroimaging in repetitive brain trauma. *Alzheimers Res Ther.* 2014;6(1):10.
34. Koerte IK, Ertl-Wagner B, Reiser M, Zafonte R, Shenton ME. White matter integrity in the brains of professional soccer players without a symptomatic concussion. *JAMA.* 2012;308(18):1859-1861.
35. Koerte IK, Lin AP, Willems A, et al. A review of neuroimaging findings in repetitive brain trauma. *Brain Pathol.* 2015;25(3):318-349.
36. Bernick C, Banks SJ, Shin W, et al. Repeated head trauma is associated with smaller thalamic volumes and slower processing speed: the Professional Fighters' Brain Health Study. *Br J Sports Med.* 2015;49(15):1007-1011.
37. Bernick C, Shan G, Zetterberg H, et al. Longitudinal change in regional brain volumes with exposure to repetitive head impacts. *Neurology.* 2020;94(3):e232-e240.
38. Singh R, Meier TB, Kuplicki R, et al. Relationship of collegiate football experience and concussion with hippocampal volume and cognitive outcomes. *Jama.* 2014;311(18):1883-1888.
39. Lepage C, Muehlmann M, Tripodis Y, et al. Limbic system structure volumes and associated neurocognitive functioning in former NFL players. *Brain Imaging Behav.* 2018.
40. Crick F, Koch C. A framework for consciousness. *Nat Neurosci.* 2003;6(2):119-126.
41. Herrero MT, Barcia C, Navarro JM. Functional anatomy of thalamus and basal ganglia. *Childs Nerv Syst.* 2002;18(8):386-404.
42. Schmahmann JD. Vascular syndromes of the thalamus. *Stroke.* 2003;34(9):2264-2278.
43. Banks SJ, Miller JB, Rissman RA, Bernick CB. Lack of Influence of Apolipoprotein E Status on Cognition or Brain Structure in Professional Fighters. *J Neurotrauma.* 2017;34(2):380-384.
44. Pellman EJ, Viano DC, Tucker AM, Casson IR, Waeckerle JF. Concussion in professional football: reconstruction of game impacts and injuries. *Neurosurgery.* 2003;53(4):799-812; discussion 812-794.

45. Viano DC, Casson IR, Pellman EJ, et al. Concussion in professional football: comparison with boxing head impacts--part 10. *Neurosurgery.* 2005;57(6):1154-1172; discussion 1154-1172.
46. Concannon LG, Kaufman MS, Herring SA. Counseling athletes on the risk of chronic traumatic encephalopathy. *Sports Health.* 2014;6(5):396-401.
47. Arciniegas DB, Bullock MR, Katz DI, Jeffrey S. Kreutzer PHDA, Ross D. Zafonte D, Zasler ND. *Brain Injury Medicine, 2nd Edition: Principles and Practice.* Springer Publishing Company; 2012.
48. Asken BM, Sullan MJ, Snyder AR, et al. Factors Influencing Clinical Correlates of Chronic Traumatic Encephalopathy (CTE): a Review. *Neuropsychol Rev.* 2016.
49. Hazrati LN, Tartaglia MC, Diamandis P, et al. Absence of chronic traumatic encephalopathy in retired football players with multiple concussions and neurological symptomatology. *Front Hum Neurosci.* 2013;7:222.
50. Alosco ML, Kasimis AB, Stamm JM, et al. Age of first exposure to American football and long-term neuropsychiatric and cognitive outcomes. *Transl Psychiatry.* 2017;7(9):e1236.
51. Mez J, Daneshvar DH, Abdolmohammadi B, et al. Duration of American Football Play and Chronic Traumatic Encephalopathy. *Ann Neurol.* 2020;87(1):116-131.
52. Stein TD, Alvarez VE, McKee AC. Concussion in Chronic Traumatic Encephalopathy. *Curr Pain Headache Rep.* 2015;19(10):47.
53. Montenigro PH, Alosco ML, Martin BM, et al. Cumulative Head Impact Exposure Predicts Later-Life Depression, Apathy, Executive Dysfunction, and Cognitive Impairment in Former High School and College Football Players. *J Neurotrauma.* 2017;34(2):328-340.
54. Asken BM, Sullan MJ, DeKosky ST, Jaffee MS, Bauer RM. Research Gaps and Controversies in Chronic Traumatic Encephalopathy: A Review. *JAMA Neurol.* 2017;74(10):1255-1262.
55. Schneider GE. Is it really better to have your brain lesion early? A revision of the "Kennard principle". *Neuropsychologia.* 1979;17(6):557-583.
56. Covassin T, Elbin RJ, Harris W, Parker T, Kontos A. The role of age and sex in symptoms, neurocognitive performance, and postural stability in athletes after concussion. *Am J Sports Med.* 2012;40(6):1303-1312.
57. Field M, Collins MW, Lovell MR, Maroon J. Does age play a role in recovery from sports-related concussion? A comparison of high school and collegiate athletes. *J Pediatr.* 2003;142(5):546-553.
58. Giza CC, Griesbach GS, Hovda DA. Experience-dependent behavioral plasticity is disturbed following traumatic injury to the immature brain. *Behav Brain Res.* 2005;157(1):11-22.
59. Harmon KG, Drezner J, Gammons M, et al. American Medical Society for Sports Medicine position statement: concussion in sport. *Clin J Sport Med.* 2013;23(1):1-18.
60. Hessen E, Nestvold K, Anderson V. Neuropsychological function 23 years after mild traumatic brain injury: a comparison of outcome after paediatric and adult head injuries. *Brain Inj.* 2007;21(9):963-979.
61. Zuckerman SL, Lee YM, Odom MJ, Solomon GS, Forbes JA, Sills AK. Recovery from sports-related concussion: Days to return to neurocognitive baseline in adolescents versus young adults. *Surg Neurol Int.* 2012;3:130.
62. Sim A, Terryberry-Spohr L, Wilson KR. Prolonged recovery of memory functioning after mild traumatic brain injury in adolescent athletes. *J Neurosurg.* 2008;108(3):511-516.
63. Beauchamp MH, Ditchfield M, Maller JJ, et al. Hippocampus, amygdala and global brain changes 10 years after childhood traumatic brain injury. *Int J Dev Neurosci.* 2011;29(2):137-143.
64. Stamm JM, Bourlas AP, Baugh CM, et al. Age of first exposure to football and later-life cognitive impairment in former NFL players. *Neurology.* 2015.

65. Stamm JM, Koerte IK, Muehlmann M, et al. Age at First Exposure to Football is Associated with Altered Corpus Callosum White Matter Microstructure in Former Professional Football Players. *J Neurotrauma.* 2015.
66. Alosco ML, Mez J, Tripodis Y, et al. Age of first exposure to tackle football and chronic traumatic encephalopathy. *Ann Neurol.* 2018;83(5):886-901.
67. Schultz V, Stern RA, Tripodis Y, et al. Age at First Exposure to Repetitive Head Impacts Is Associated with Smaller Thalamic Volumes in Former Professional American Football Players. *J Neurotrauma.* 2018;35(2):278-285.
68. Stamm JM, Bourlas AP, Baugh CM, et al. Age of first exposure to football and later-life cognitive impairment in former NFL players. *Neurology.* 2015;84(11):1114-1120.
69. Baugh CM, Kiernan PT, Kroshus E, et al. Frequency of head-impact-related outcomes by position in NCAA division I collegiate football players. *J Neurotrauma.* 2015;32(5):314-326.
70. Stein TD, Montenigro PH, Alvarez VE, et al. Beta-amyloid deposition in chronic traumatic encephalopathy. *Acta Neuropathol.* 2015;130(1):21-34.
71. McCrory P, Meeuwisse WH, Kutcher JS, Jordan BD, Gardner A. What is the evidence for chronic concussion-related changes in retired athletes: behavioural, pathological and clinical outcomes? *Br J Sports Med.* 2013;47(5):327-330.
72. Watanabe T, Elovic E, Zafonte R. Chronic traumatic encephalopathy. *Pm r.* 2010;2(7):671-675.
73. Deng H, Ordaz A, Upadhyayula PS, et al. Apolipoprotein E Epsilon 4 Genotype, Mild Traumatic Brain Injury, and the Development of Chronic Traumatic Encephalopathy. *Med Sci (Basel).* 2018;6(3).
74. Omalu BI, Fitzsimmons RP, Hammers J, Bailes J. Chronic traumatic encephalopathy in a professional American wrestler. *J Forensic Nurs.* 2010;6(3):130-136.
75. Stern RA, Riley DO, Daneshvar DH, Nowinski CJ, Cantu RC, McKee AC. Long-term consequences of repetitive brain trauma: chronic traumatic encephalopathy. *PM R.* 2011;3(10 Suppl 2):S460-467.
76. Shim SS, Stutzmann GE. Inhibition of Glycogen Synthase Kinase-3: An Emerging Target in the Treatment of Traumatic Brain Injury. *J Neurotrauma.* 2016;33(23):2065-2076.
77. Huettel SA, Song AW, McCarthy G. *Functional Magnetic Resonance Imaging.* Oxford University Press, Incorporated; 2009.
78. Grover VP, Tognarelli JM, Crossey MM, Cox IJ, Taylor-Robinson SD, McPhail MJ. Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians. *J Clin Exp Hepatol.* 2015;5(3):246-255.
79. Felix R, Langer R. *Klinische Radiologie mit Repetitorium.* De Gruyter; 2011.
80. Banks SJ, Mayer B, Obuchowski N, et al. Impulsiveness in professional fighters. *J Neuropsychiatry Clin Neurosci.* 2014;26(1):44-50.
81. Fischl B, Salat DH, Busa E, et al. Whole Brain Segmentation. *Neuron.* 2002;33(3):341-355.
82. Grimm O, Pohlack S, Cacciaglia R, et al. Amygdalar and hippocampal volume: A comparison between manual segmentation, Freesurfer and VBM. *J Neurosci Methods.* 2015;253:254-261.
83. Fischl B. FreeSurfer. *Neuroimage.* 2012;62(2):774-781.
84. Fischl B, van der Kouwe A, Destrieux C, et al. Automatically parcellating the human cerebral cortex. *Cereb Cortex.* 2004;14(1):11-22.
85. Reuter M, Rosas HD, Fischl B. Highly accurate inverse consistent registration: a robust approach. *Neuroimage.* 2010;53(4):1181-1196.
86. Segonne F, Dale AM, Busa E, et al. A hybrid approach to the skull stripping problem in MRI. *Neuroimage.* 2004;22(3):1060-1075.
87. Talairach J, Tournoux P. *Co-planar Stereotaxic Atlas of the Human Brain: 3-dimensional Proportional System : an Approach to Cerebral Imaging.* G. Thieme; 1988.

88. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging.* 1998;17(1):87-97.
89. Fischl B, Salat DH, van der Kouwe AJ, et al. Sequence-independent segmentation of magnetic resonance images. *Neuroimage.* 2004;23 Suppl 1:S69-84.
90. Alosco ML, Koerte IK, Tripodis Y, et al. White matter signal abnormalities in former National Football League players. *Alzheimers Dement (Amst).* 2018;10:56-65.
91. Lepage C, Muehlmann M, Tripodis Y, et al. Limbic system structure volumes and associated neurocognitive functioning in former NFL players. *Brain Imaging Behav.* 2019;13(3):725-734.
92. Wojtowicz M, Gardner AJ, Stanwell P, Zafonte R, Dickerson BC, Iverson GL. Cortical thickness and subcortical brain volumes in professional rugby league players. *Neuroimage Clin.* 2018;18:377-381.
93. Hart MG, Housden CR, Suckling J, et al. Advanced magnetic resonance imaging and neuropsychological assessment for detecting brain injury in a prospective cohort of university amateur boxers. *Neuroimage Clin.* 2017;15:194-199.
94. Drijkoningen D, Chalavi S, Sunaert S, Duysens J, Swinnen SP, Caeyenberghs K. Regional Gray Matter Volume Loss Is Associated with Gait Impairments in Young Brain-Injured Individuals. *J Neurotrauma.* 2017;34(5):1022-1034.
95. Koerte IK, Mayinger M, Muehlmann M, et al. Cortical thinning in former professional soccer players. *Brain Imaging Behav.* 2015.
96. Iscan Z, Jin TB, Kendrick A, et al. Test-retest reliability of freesurfer measurements within and between sites: Effects of visual approval process. *Hum Brain Mapp.* 2015;36(9):3472-3485.
97. Qiu A, Brown T, Fischl B, Ma J, Miller MI. Atlas generation for subcortical and ventricular structures with its applications in shape analysis. *IEEE Trans Image Process.* 2010;19(6):1539-1547.
98. Akudjedu TN, Nabulsi L, Makelyte M, et al. A comparative study of segmentation techniques for the quantification of brain subcortical volume. *Brain Imaging Behav.* 2018;12(6):1678-1695.
99. Alexander B, Georgiou-Karistianis N, Beare R, et al. Accuracy of automated amygdala MRI segmentation approaches in Huntington's disease in the IMAGE-HD cohort. *Hum Brain Mapp.* 2020;41(7):1875-1888.
100. Arnold SJ, Ivleva EI, Gopal TA, et al. Hippocampal volume is reduced in schizophrenia and schizoaffective disorder but not in psychotic bipolar I disorder demonstrated by both manual tracing and automated parcellation (FreeSurfer). *Schizophr Bull.* 2015;41(1):233-249.
101. Bartel F, Vrenken H, Bijma F, Barkhof F, van Herk M, de Munck JC. Regional analysis of volumes and reproducibilities of automatic and manual hippocampal segmentations. *PLoS One.* 2017;12(2):e0166785.
102. Beelen C, Phan TV, Wouters J, Ghesquiere P, Vandermosten M. Investigating the Added Value of FreeSurfer's Manual Editing Procedure for the Study of the Reading Network in a Pediatric Population. *Front Hum Neurosci.* 2020;14:143.
103. Brinkmann BH, Guragain H, Kenney-Jung D, et al. Segmentation errors and intertest reliability in automated and manually traced hippocampal volumes. *Ann Clin Transl Neurol.* 2019;6(9):1807-1814.
104. Cherbuin N, Anstey KJ, Reglade-Meslin C, Sachdev PS. In vivo hippocampal measurement and memory: a comparison of manual tracing and automated segmentation in a large community-based sample. *PLoS One.* 2009;4(4):e5265.
105. Clerx L, Gronenschild EH, Echavarri C, Verhey F, Aalten P, Jacobs HI. Can FreeSurfer Compete with Manual Volumetric Measurements in Alzheimer's Disease? *Curr Alzheimer Res.* 2015;12(4):358-367.

106. de Flores R, La Joie R, Landeau B, et al. Effects of age and Alzheimer's disease on hippocampal subfields: comparison between manual and FreeSurfer volumetry. *Hum Brain Mapp.* 2015;36(2):463-474.
107. Dewey J, Hana G, Russell T, et al. Reliability and validity of MRI-based automated volumetry software relative to auto-assisted manual measurement of subcortical structures in HIV-infected patients from a multisite study. *Neuroimage.* 2010;51(4):1334-1344.
108. Doring TM, Kubo TT, Cruz LC, Jr., et al. Evaluation of hippocampal volume based on MR imaging in patients with bipolar affective disorder applying manual and automatic segmentation techniques. *J Magn Reson Imaging.* 2011;33(3):565-572.
109. Fraser MA, Shaw ME, Anstey KJ, Cherbuin N. Longitudinal Assessment of Hippocampal Atrophy in Midlife and Early Old Age: Contrasting Manual Tracing and Semi-automated Segmentation (FreeSurfer). *Brain Topogr.* 2018;31(6):949-962.
110. Fung YL, Ng KET, Vogrin SJ, et al. Comparative Utility of Manual versus Automated Segmentation of Hippocampus and Entorhinal Cortex Volumes in a Memory Clinic Sample. *J Alzheimers Dis.* 2019;68(1):159-171.
111. Germeyan SC, Kalikhman D, Jones L, Theodore WH. Automated versus manual hippocampal segmentation in preoperative and postoperative patients with epilepsy. *Epilepsia.* 2014;55(9):1374-1379.
112. Herten A, Konrad K, Krinzinger H, Seitz J, von Polier GG. Accuracy and bias of automatic hippocampal segmentation in children and adolescents. *Brain Struct Funct.* 2019;224(2):795-810.
113. Hosseini MP, Nazem-Zadeh MR, Pompili D, Jafari-Khouzani K, Elisevich K, Soltanian-Zadeh H. Comparative performance evaluation of automated segmentation methods of hippocampus from magnetic resonance images of temporal lobe epilepsy patients. *Med Phys.* 2016;43(1):538.
114. Keller SS, Gerdes JS, Mohammadi S, et al. Volume estimation of the thalamus using freesurfer and stereology: consistency between methods. *Neuroinformatics.* 2012;10(4):341-350.
115. Kempton MJ, Underwood TS, Brunton S, et al. A comprehensive testing protocol for MRI neuroanatomical segmentation techniques: Evaluation of a novel lateral ventricle segmentation method. *Neuroimage.* 2011;58(4):1051-1059.
116. Khlif MS, Egorova N, Werden E, et al. A comparison of automated segmentation and manual tracing in estimating hippocampal volume in ischemic stroke and healthy control participants. *Neuroimage Clin.* 2019;21:101581.
117. Kim H, Chupin M, Colliot O, Bernhardt BC, Bernasconi N, Bernasconi A. Automatic hippocampal segmentation in temporal lobe epilepsy: impact of developmental abnormalities. *Neuroimage.* 2012;59(4):3178-3186.
118. Lehmann M, Douiri A, Kim LG, et al. Atrophy patterns in Alzheimer's disease and semantic dementia: a comparison of FreeSurfer and manual volumetric measurements. *Neuroimage.* 2010;49(3):2264-2274.
119. Makowski C, Beland S, Kostopoulos P, et al. Evaluating accuracy of striatal, pallidal, and thalamic segmentation methods: Comparing automated approaches to manual delineation. *Neuroimage.* 2018;170:182-198.
120. Morey RA, Petty CM, Xu Y, et al. A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *Neuroimage.* 2009;45(3):855-866.
121. Mulder ER, de Jong RA, Knol DL, et al. Hippocampal volume change measurement: quantitative assessment of the reproducibility of expert manual outlining and the automated methods FreeSurfer and FIRST. *Neuroimage.* 2014;92:169-181.

122. Olsson E, Klasson N, Berge J, et al. White Matter Lesion Assessment in Patients with Cognitive Impairment and Healthy Controls: Reliability Comparisons between Visual Rating, a Manual, and an Automatic Volumetrical MRI Method-The Gothenburg MCI Study. *J Aging Res.* 2013;2013:198471.
123. Pardoe HR, Pell GS, Abbott DF, Jackson GD. Hippocampal volume assessment in temporal lobe epilepsy: How good is automated segmentation? *Epilepsia.* 2009;50(12):2586-2592.
124. Perlaki G, Horvath R, Nagy SA, et al. Comparison of accuracy between FSL's FIRST and Freesurfer for caudate nucleus and putamen segmentation. *Sci Rep.* 2017;7(1):2418.
125. Ranta ME, Chen M, Crocetti D, et al. Automated MRI parcellation of the frontal lobe. *Hum Brain Mapp.* 2014;35(5):2009-2026.
126. Sanchez-Benavides G, Gomez-Anson B, Sainz A, Vives Y, Delfino M, Pena-Casanova J. Manual validation of FreeSurfer's automated hippocampal segmentation in normal aging, mild cognitive impairment, and Alzheimer Disease subjects. *Psychiatry Res.* 2010;181(3):219-225.
127. Schmidt MF, Storrs JM, Freeman KB, et al. A comparison of manual tracing and FreeSurfer for estimating hippocampal volume over the adult lifespan. *Hum Brain Mapp.* 2018;39(6):2500-2513.
128. Schoemaker D, Buss C, Head K, et al. Hippocampus and amygdala volumes from magnetic resonance images in children: Assessing accuracy of FreeSurfer and FSL against manual segmentation. *Neuroimage.* 2016;129:1-14.
129. Shen L, Saykin AJ, Kim S, et al. Comparison of manual and automated determination of hippocampal volumes in MCI and early AD. *Brain Imaging Behav.* 2010;4(1):86-95.
130. Stenkrona P, Matheson GJ, Cervenka S, Sigray PP, Halldin C, Farde L. [(11)C]SCH23390 binding to the D1-dopamine receptor in the human brain-a comparison of manual and automated methods for image analysis. *EJNMMI Res.* 2018;8(1):74.
131. Tae WS, Kim SS, Lee KU, Nam EC, Kim KW. Validation of hippocampal volumes measured using a manual method and two automated methods (FreeSurfer and IBASPM) in chronic major depressive disorder. *Neuroradiology.* 2008;50(7):569-581.
132. Tavares TP, Mitchell DGV, Coleman K, et al. Ventricular volume expansion in presymptomatic genetic frontotemporal dementia. *Neurology.* 2019;93(18):e1699-e1706.
133. Wenger E, Martensson J, Noack H, et al. Comparing manual and automatic segmentation of hippocampal volumes: reliability and validity issues in younger and older brains. *Hum Brain Mapp.* 2014;35(8):4236-4248.
134. Alosco ML, Jarnagin J, Tripodis Y, et al. Olfactory Function and Associated Clinical Correlates in Former National Football League Players. *J Neurotrauma.* 2016.
135. Marin RSB, R.C.; Firinciogullari, S. . *Reliability and validity of the Apathy Evaluation Scale.* Psychiatry Res 38; 1991.
136. Beck AS, R.; Brown, G. *Manual for the BDI-II.* San Antonio, TX: Psychological Corporation; 1996.
137. Beck AS, R. . *Manual for the Beck hopelessness scale.* San Antonio, TX: Psychological Corporation; 1988.
138. Barratt ES. Factor analysis of some psychometric measures of impulsiveness and anxiety. *Psychol Rep.* 1965;16:547-554.
139. Roth RM, Isquith PK, Gioia GA. *BRIEF-A : behavior rating inventory of executive function--adult version : professional manual.* Lutz, Florida: Psychological Assessment Resources; 2005.
140. Radloff LS. *The CES-D scale a self-report depression scale for research in the general population.* Vol 1: Applied psychological measurement; 1977.
141. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry.* 1960;23:56-61.

142. Brown GLG, F.K.; Ballenger, J.C.; Goyer, P.F.; Major, L.F. *Aggression in humans correlates with cerebrospinal fluid amine metabolites.* Psychiatry research; 1979.
143. Miller IWN, W.H.; Bishop, S.B.; Dow, M.G. *The Modified Scale for Suicidal Ideation: reliability and validity.* Vol 54: J Consult Clin Psychol; 1986.
144. Buss AH, Durkee A. An inventory for assessing different kinds of hostility. *J Consult Psychol.* 1957;21(4):343-349.
145. Lezak MDH, D.B.; Bigler, E. D.; Tranel, D. . *Neuropsychological assessment.* 5th ed, Oxford: Oxford University Press. xxv; 2012.
146. Delis DC, Kaplan E, Kramer JH. Delis Kaplan Executive Function System (D-KEFS). Vol The Psychological Corporation. San Antonio, TX2001.
147. Reitan RM. *Trail Making Test: Manual for administration and scoring: Reitan Neuropsychology Laboratory.* Tucson, AZ1992.
148. Wechsler D. Manual for the Wechsler Adult Intelligence Scale-III. San Antonio, TX: Psychological Corporation; 1997.
149. Stern R, White T. *The Neuropsychological Assessment Battery (NAB).* Lutz, Florida: Psychological Assessment Resources; 2003.
150. Stern RA, Javorsky DJ, Singer EA, et al. *Boston Qualitative Scoring System (BQSS) for the Rey-Osterrieth complex figure.* Odessa, FL: Psychological Assessment Resources (PAR); 1999.
151. Guenette JP, Stern RA, Tripodis Y, et al. Automated versus manual segmentation of brain region volumes in former football players. *Neuroimage Clin.* 2018;18:888-896.
152. Riffkin J, Yucel M, Maruff P, et al. A manual and automated MRI study of anterior cingulate and orbito-frontal cortices, and caudate nucleus in obsessive-compulsive disorder: comparison with healthy controls and patients with schizophrenia. *Psychiatry Res.* 2005;138(2):99-113.
153. Mueller SG, Yushkevich PA, Das S, et al. Systematic comparison of different techniques to measure hippocampal subfield volumes in ADNI2. *Neuroimage Clin.* 2018;17:1006-1018.
154. Fischbach-Boulanger C, Fitsiori A, Noblet V, et al. T1- or T2-weighted magnetic resonance imaging: what is the best choice to evaluate atrophy of the hippocampus? *Eur J Neurol.* 2018;25(5):775-781.
155. Phelps EA. Human emotion and memory: interactions of the amygdala and hippocampal complex. *Curr Opin Neurobiol.* 2004;14(2):198-202.
156. Zagorchev L, Meyer C, Stehle T, et al. Differences in Regional Brain Volumes Two Months and One Year after Mild Traumatic Brain Injury. *J Neurotrauma.* 2016;33(1):29-34.
157. Chen XY, Wang Q, Wang X, Wong KS. Clinical Features of Thalamic Stroke. *Curr Treat Options Neurol.* 2017;19(2):5.
158. Zhao YJ, Du MY, Huang XQ, et al. Brain grey matter abnormalities in medication-free patients with major depressive disorder: a meta-analysis. *Psychol Med.* 2014;44(14):2927-2937.
159. Konstantinou N, Pettemeridou E, Seimenis I, et al. Assessing the Relationship between Neurocognitive Performance and Brain Volume in Chronic Moderate-Severe Traumatic Brain Injury. *Front Neurol.* 2016;7:29.
160. Robinson S, Berglass JB, Denson JL, et al. Microstructural and microglial changes after repetitive mild traumatic brain injury in mice. *J Neurosci Res.* 2017;95(4):1025-1035.
161. Okada N, Fukunaga M, Yamashita F, et al. Abnormal asymmetries in subcortical brain volume in schizophrenia. *Mol Psychiatry.* 2016.
162. Eshaghi A, Marinescu RV, Young AL, et al. Progression of regional grey matter atrophy in multiple sclerosis. *Brain.* 2018;141(6):1665-1677.
163. Grodin EN, Momenan R. Decreased subcortical volumes in alcohol dependent individuals: effect of polysubstance use disorder. *Addict Biol.* 2016.

164. Stern RA, Adler CH, Chen K, et al. Tau Positron-Emission Tomography in Former National Football League Players. *N Engl J Med.* 2019;380(18):1716-1725.
165. Koerte IK, Schultz V, Sydnor VJ, et al. Sex-Related Differences in the Effects of Sports-Related Concussion: A Review. *J Neuroimaging.* 2020;30(4):387-409.