**OPEN**

Correspondence and
requests for materials
should be addressed to
A.H.
(adamdghampshire@
gmail.com)

# Hypoconnectivity and Hyperfrontality in Retired American Football Players

Adam Hampshire[1,2], Alex MacDonald[2] & Adrian M. Owen[2]

[1]The Computational, Cognitive and Clinical Neuroimaging Laboratory, The Division of Brain Sciences, Imperial College London, UK, [2]The Brain and Mind Institute, The Natural Sciences Centre, University of Western Ontario, London, Ontario, Canada.

**Recent research has raised concerns about the long-term neurological consequences of repetitive concussive and sub-concussive injuries in professional players of American Football. Despite this interest, the neural and psychological status of retired players remains unknown. Here, we evaluated the performances and brain activation patterns of retired National Football League players (NFL alumni) relative to controls using an fMRI-optimised neuropsychological test of executive function. Behaviourally, the NFL alumni showed only modest performance deficits on the executive task. By contrast, they showed pronounced hyperactivation and hypoconnectivity of the dorsolateral frontal and frontopolar cortices. Critically, abnormal frontal-lobe function was correlated with the number of times that NFL alumni reported having been removed from play after head injury and was evident in individual players. These results support the hypothesis that NFL alumni have a heightened probability of developing executive dysfunction and suggest that fMRI provides the most sensitive biomarker of the underlying neural abnormality.**

Recent research has raised concerns regarding the prevalence of mild traumatic brain injuries (mTBI) and the potential for associated long-term neurological damage in players of contact sports like American professional football[1]. It is estimated based on statistics from the Centers for Disease Control and Prevention that between 1.6 and 3.8 million sports concussions occur annually in the United States alone[2]. Furthermore, research has demonstrated that American football players can experience hundreds of sub-concussive hits during a single season[3–8]. While the long-term consequences of repeated mTBI remains unclear, recent investigations indicate that NFL alumni have a heightened risk of developing a host of debilitating disorders[9,10]. For example, a report commissioned by the National Football League (NFL), showed that retired players between the ages of 30–49 were 20 times more likely, at a rate of 1.9%, to receive a diagnosis of dementia, Alzheimer's disease (AD), or other memory-related impairments[11]. Over the age of 50, the proportion diagnosed with one of the above rose to 6.1%, compared with only 1.2% In the general population[11]. Overall, neurodegenerative mortality of NFL alumni is 3 times greater than that of the general population, with AD and amyotrophic lateral sclerosis (ALS) estimated at 4 times as high[10]. Finally, NFL alumni may be at increased risk of developing chronic traumatic encephalopathy (CTE), a progressive tauopathy that is characterized by behavioural changes, memory impairments and Parkinsonism[12,13]. In fact, one recent publication reported that under post-mortem analyses 89% of a non-random sample of NFL alumni showed evidence of CTE[14]. Importantly, to date, no reliable method has been established for detecting and tracking the long-term consequences of repetitive mTBI in vivo. Consequently, the long-term impact of repeated mTBI on brain function remains to be determined.

Indeed, despite these alarming statistics, neuropsychological data regarding the long-term effects of mTBI on NFL players are contradictory. For example, whereas one recent study[15] reported that NFL alumni were impaired on global measures of cognitive function relative to controls, earlier investigations failed to find persisting decrements in neuropsychological performance in active players post-injury[7], a result that is characteristic of the sports mTBI literature. More broadly, a discrepancy between real-life outcomes and performance on neuropsychological tests of executive function is common for individuals who suffer mild to moderate TBI[16]. One possible explanation for this discrepancy is that lab based executive neuropsychological tests lack ecological validity. More specifically, compensatory cortical mechanisms might be sufficient to counter the behavioural consequences of neurological damage under controlled laboratory conditions but insufficient in complex real world scenarios[17,18]. While NFL alumni are a particularly select and therefore hard to recruit population, the number of head collisions that they suffer of great enough severity to warrant being removed from play varies greatly between players. Consequently, they offer a unique opportunity to examine the way in which individual differences in the number of mTBIs suffered relate to individual differences in executive cognition and to thereby derive broadly applicable insights into how long-term cortical plasticity may work to counter neurological damage.

Here, we assessed the performances and functional brain activations of a group of retired NFL players whilst undertaking the One Touch Spatial Planning task[19], an fMRI-optimised variant of the classical Tower of London paradigm that is widely used to assess executive cognition[20]. It was hypothesized that, consistent with a recent meta-analysis of the effects of multiple concussions[21], NFL alumni would show only a modest behavioural impairment. Critically, it was also predicted that fMRI would reveal more robust evidence of long-term cortical compensatory mechanisms, with abnormalities in task-related functional brain activity and connectivity.

## Results

**Behavioural analysis.** The One Touch Spatial Planning task (SP – Fig. 1a) is an fMRI-optimised variant on the Cambridge Neuropsychological Test Automated Battery (CANTAB) task of the same name, which in turn is based on the popular Tower of London paradigm[20]. In order to assess behaviour, the total numbers of correctly solved planning problems were examined with a 2 * 3 mixed ANOVA in SPSS, in which the within-subject factor was 2, 3 or 4 move problems (Planning Complexity) and the between-subject factor was NFL alumni vs. Control (Group). There was a significant main effect of Planning Complexity ($F_{2,62} = 6.176$ p =

0.014), no significant main effect of Group ($F_{1,31} = 2.462$ p = 0.127), and a sub-threshold Group * Planning Complexity interaction ($F_{2,62} = 3.04$ p = 0.055). Comparing total number of problems solved separately for 2, 3 and 4 move problems showed a weak effect of Group for the 4 move problems, with NFL alumni solving fewer problems than controls (2 move t = 0.073 p = 0.471, 3 move t = 1.055 p = 0.150, 4 move t = 2.323 p = 0.014, values 1-tailed and uncorrected). There were no significant non-parametric correlations between total number of problems solved and the ranked number of times that the NFL players reported having been sent off field due to head impact (RTO – mean = 6.07 SD = 8.24). When comparing the total number of correctly solved problems on an individual participant level, 2 out of 13 NFL alumni were ranked at or below the bottom 10th percentile relative to controls. Examining response times in a factorial analysis of the same design showed a main effect of Planning Complexity ($F_{2,62} = 17.690$ p < 0.001), with no significant main effect of group ($F_{1,31} = 1.742$ p = 0.197) and no significant interaction with Group ($F_{2,62} = 1.286$ p = 0.284). There was no significant non-parametric correlation between mean RT during planning and RTO. When comparing overall mean RT individually, 4 out of the 13 NFL alumni were ranked at or above the slowest 10th percentile relative to controls.
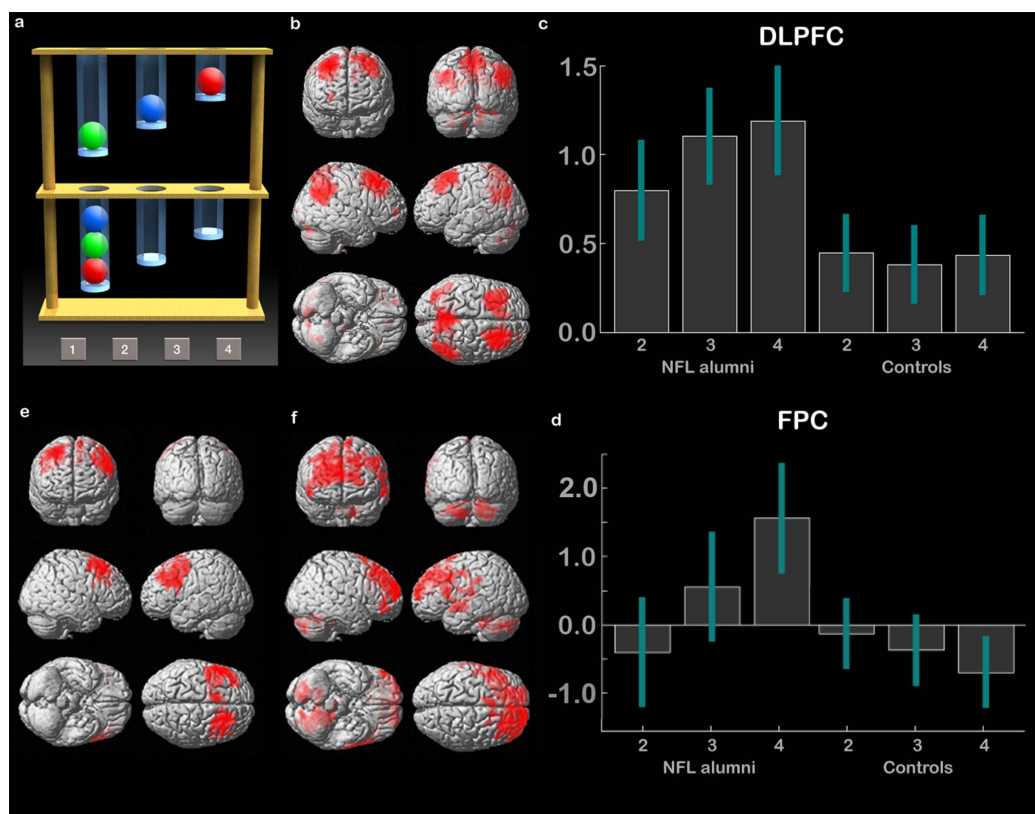


**Figure 1 | The Spatial Planning Task (SP).** (*1a*) The SP task is a variant on the Tower of London paradigm[20], which is commonly used to assess executive function. Two arrays of coloured balls are presented in vertical tubes. In the planning condition, the participant must work out the minimum number of moves required to rearrange the balls in the top array so that they match those in the bottom array. Difficulty varies from 2 to 4 moves. There are also counting trials, in which different numbers of balls are placed in the two arrays and the participant must subtract the number of balls at the top from the number at the bottom. (*1b*) In the fMRI analysis, the counting condition is used as a visual and attentional control. Contrasting planning minus counting (voxel-wise FDR correction for the whole brain mass p < 0·05) localises a dorsal frontoparietal network including middle frontal gyrus, frontopolar cortex, and inferior parietal cortex[32]. (*1c*) In the ROI analysis, there was significantly greater activation in the DLPFC in NFL alumni relative to controls, particularly at the more difficult planning levels (error bars report 90% confidence interval). (*1d*) There was a significant interaction between planning level and group within the FPC ROI (error bars report 90% confidence interval). (*1e*) In accordance with the ROI analysis, NFL alumni showed extensive clusters of hyperactivation within the DLPFC bilaterally (rendered at p < 0·01 with 500 voxel extent threshold. Peak coordinates significant with whole brain FDR voxel-wise correction – **see table 1a**). (*1f*) The interaction between Group and Planning Level showed extensive clusters of hyperactivation in NFL alumni relative to controls within the frontopolar cortices (rendered at p < 0·01 uncorrected with a 500 voxel extent threshold. Peak activation coordinates significant with whole brain FDR voxel-wise correction – **see table 1b**).

Table 1 | peak coordinates from analysis of the Spatial Planning Task

| | X | Y | Z | t |
|---|---|---|---|---|
| **A) Planning NFL > Controls** | | | | |
| Left DLPFC | −48 | 2 | 56 | 5·27 |
| Left DLPFC | −48 | 28 | 42 | 4·94 |
| Right DLPFC | 28 | 22 | 64 | 4·52 |
| Right DLPFC | 32 | 26 | 58 | 4·26 |
| preSMA | −8 | 10 | 52 | 4·31 |
| **B) Effect of planning level NFL > controls** | | | | |
| Left FPC | 8 | 62 | 22 | 4·97 |
| Right FPC | −12 | 60 | 14 | 4·91 |
| ACC | 4 | 46 | 24 | 3·89 |
| Left cerebellum | −18 | −58 | −44 | 4·18 |
| Right cerebellum | 14 | −78 | −42 | 3·87 |
| **C) Non-parametric effect of RTO** | | | | |
| RDLPFC | 42 | 26 | 42 | 4·20 |
| RDLPFC | 42 | 4 | 48 | 3·78 |
| RFPC | 48 | 42 | 22 | 3·99 |
| Precuneous | 4 | −68 | 58 | 5·24 |
| RPC | 32 | −66 | 54 | 4·60 |

**Functional ROI analysis.** In the neuroimaging analysis, contrasting activation during planning minus activation during counting, collapsed across NFL alumni and controls, localised the expected dorsal frontoparietal network including clusters within dorsolateral prefrontal cortex (DLPFC), parietal cortex (PC) and frontopolar cortex (FPC). (**Fig. 1b**). These clusters were used to form 3 * bilateral ROIs from which mean activations were extracted for further analysis. The resultant data (**Fig. 1c**) were analysed using three full factorial models, in which the within-subject factor was activation relative to rest for 2, 3 and 4 move problems (Planning Complexity) and the between-subject factor was NFL alumni vs. control (Group). There was a significant main effect of Group in the DLPFC ($F_{1,31}$ = 8.72 P = 0.004), but not within the PC ($F_{1,31}$ = 0.68 P = 0.409) or FPC ($F_{1,31}$ = 0.28 p = 0.597). There was also a significant main effect of Planning Complexity within the DLPFC ($F_{2,62}$ = 7.02 p = 0.009), but not within the PC ($F_{2,62}$ = 0.59 p = 0.443) or FPC ($F_{2,62}$ = 1.60 p = 0.207). There was a significant interaction of Group * Planning Complexity within the DLPFC ($F_{2,62}$ = 7.88 p = 0.005) and FPC (**Fig. 1d**) ($F_{2,62}$ = 7.34 P = 0.007), but not within the PC ($F_{2,62}$ = 0.52 p = 0.470). These results were driven by hyperactivation in NFL alumni within the DLPFC at all levels of planning complexity and within the FPC selectively at the higher levels of planning demand. On an individual participant level, 7 out of 13 NFL alumni were at or above the top 10 percentile relative to controls in terms of mean DLPFC activation during planning.

**Functional whole brain analysis.** Supplemental whole brain analyses were carried out using a 2 * 3 full factorial model in which the within-subject factor was Group and the between-subject factor was Planning Complexity. In close concordance with the ROI analyses, there was a main effect of group, with NFL alumni showing hyperactivation in bilateral DLPFC clusters (**Fig. 1e Table 1a**). More focal clusters were evident within PC bilaterally. There was a significant interaction, with additional recruitment of FPC in NFL alumni during performance of the most complex planning problems (**Fig. 1f Table 1b**). Non-parametric analyses, using RTO as the predictor, rendered an extensive cluster of voxels within the right DLPFC and PC (**Fig. 2a Table 1c**). No significant negative correlations were evident at the cluster corrected threshold.

**Functional connectivity analysis.** Functional connectivity within the dorsal planning network was examined using a psychophysiological interaction (PPI) model in SPM8. In a PPI model, the timecourse of activation is extracted from a 'seed' region (the physiological timecourse) and is interacted with a timecourse that captures different psychological conditions. Here, the physiological timecourse was extracted from a 10 mm sphere based at the peak activation foci for planning vs. counting in the right DLPFC (x = 28 y = 32 z = 52). The psychological timecourse was defined by time engagement in planning trials as +1, time engagement in counting trials as −1, and rest periods as 0. The psychological, physiological, and PPI timecourses were used to predict activation elsewhere in the
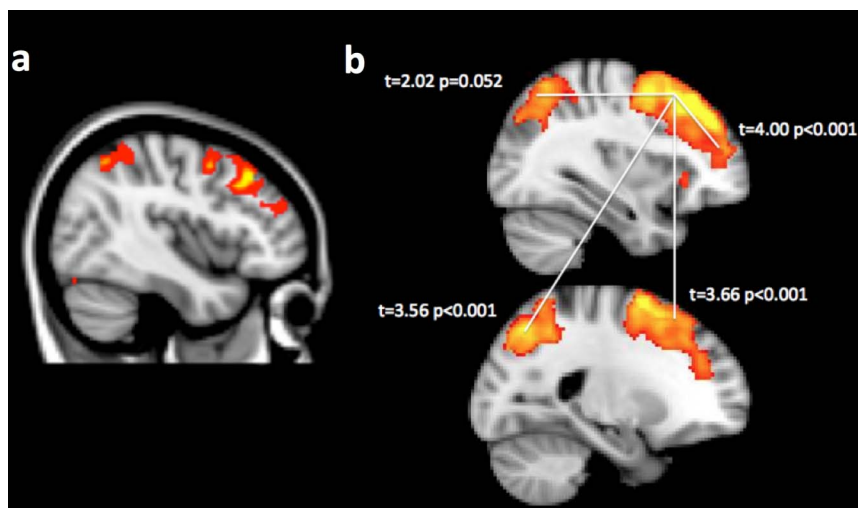


**Figure 2 | Non-parametric correlations and functional connectivity analysis.** (2a) When the ranked number of times NFL alumni had been sent off the field due to head injury (Ranked Times Off - RTO) was used to predict activation during the SP task in a non-parametric permutation analysis, extensive clusters of activation were evident in the right DLPFC extending into the FPC. A more focal cluster was also evident within the PC (image rendered at p < 0·01 uncorrected with FWE cluster correction within a mask that includes all regions that are active during task relative to rest – see table 1c). (2b) When beta weights for the physiological predictor function were contrasted across groups, NFL alumni showed generally lowered functional connectivity between the right DLPFC seed region and the left DLPFC, the right FPC and the PC bilaterally. Yellow regions show the t-test of beta weights for the physiological predictor collapsed across participants with FWE voxel-wise correction of p < 0·05 for the whole brain mass t and p values report cross-group analyses for the indicated connections.

brain at the single subject level using general linear modelling. The latter two predictor-functions generate distinct measures of functional connectivity between brain regions. Specifically, the physiological function captures the *general* correlation in activation between brain regions once activation due to the specified psychological conditions has been accounted for. Conversely, the PPI captures *psychologically specific* correlation – that is, the difference in the correlations between brain regions when the specified psychological conditions vary. Mean beta weights for the psychological and PPI predictors were extracted from 10 mm spherical ROIs defined at peak activation coordinates for the contrast of planning vs. counting trials (left DLPFC x = −22 y = 14 z = 58; left PC x = −34 y = −78 z = 38; right PC x = 44 y = −70 z = 34; right FPC x = 28 y = 50 z = 4) and the resultant data analysed at the group level. Examining the beta weights for the PPI predictor functions collapsed across groups in a one-sample t-test did not reveal any significant condition dependent correlations in activation (**Table 2a**). Nor were there any significant cross-group differences in the strength of the PPI (**Table 2b**). Examining the beta weights for the physiological predictor function collapsed across groups in a one-sample t-test revealed strong connectivity between the right DLPFC and the other dorsal planning network ROIs (**Table 2c – Fig. 2b**). Contrasting the data across groups revealed generally lowered functional connectivity for NFL alumni relative to controls in all ROIs (**Table 2D – Fig. 2B**). 12/13 NFL alumni were at or below the bottom 10% of controls in terms of mean functional connectivity between right DLPFC and other sub-regions of the dorsal planning network. There was a strong negative correlation (r = −0.61) between RTO and mean functional connectivity between right DLPFC and other dorsal planning network sub-regions.

**Affective connectivity analysis.** Whereas functional connectivity examines correlations between different brain regions at the same time point, affective connectivity measures the level at which activation in one region is predicted by activations that occurred at earlier time points elsewhere in the brain. In this respect, affective connectivity gives a measure of the activation flow between network sub-components. Here we used Granger Causality (GC) to examine

whether there were differences between NFL alumni and controls in the affective connectivity of the dorsal frontoparietal network. Mean activation timecourses were extracted from 10 mm spherical ROIs defined at peak coordinates for the contrast of planning vs. counting within the right DLPFC, right FPC and right PPC (see previous section). These timecourses were pre-processed and a granger causality model fitted to each individual's data set using the Granger Causality Connectivity Analysis Toolbox[22,23] (see methods). When GC measures were averaged across all connections and compared across groups (**Fig. 3 & Table 3**), NFL alumni showed significantly lowered mean GC relative to controls. Comparing GC measures across groups for each connection (**Table 3**), revealed significantly lowered GC in NFL alumni relative to controls in connections to the DLPFC from the PC and the FPC, and in connections to the FPC from the PC. Three of the NFL alumni were within the bottom 10% relative to controls in terms mean GC. There were no significant correlation between RTO and GC.

**Development of an mTBI polymarker using machine learning.** A support vector machine (SVM) was used to generate a polymarker for classifying NFL alumni and distance from the hyperplane was calculated to examine the relationship between classification dimensions and the number of mTBIs suffered. Data from the analyses reported in the previous sections were arranged into an input matrix that included behavioural performance measures at each level of planning demand (2, 3 and 4 move problems), activation levels for each ROI (DLPFC, PC, FPC bilaterally) at each level of planning demand, psychophysiological and physiological weights from each ROI from the PPI analysis, and the GC connectivity weights. Each row in the input matrix contained all measures from one individual and each column contained the same measure for all individuals (e.g. performance at level m4). There was also a column with the categorical (NFL vs. Control) labels. A leave-one-out validation approach was applied as follows. One participant was removed and the SVM identified the most optimal hyperplane for classifying the remaining individuals as NFL alumni or controls based on the combination of imaging and behavioural measures. The trained machine then attempted to classify the remaining individual, to whose data it was naïve. This process was repeated until each individual had been excluded and classified once. The entire process was then repeated 10,000 times, but with the categorical labels randomly permuted in order to generate a null distribution (mean accuracy = 49.3% SD = +/−11.1%). When all of the behavioural and neuroimaging measures were input to the SVM together, it classified individuals with an overall accuracy of ~84% (p < 0.001, accuracy for NFL alumni = 77% accuracy for controls = 89%). When the SVM machine was trained using all data sets together, distance from the hyperplane correlated significantly with the RTO measure in the NFL alumni group (non-parametric correlation = 0.55 p < 0.05). Further exploratory analyses were conducted to gauge the accuracy of categorisation when using data from individual analyses, these being the behavioural performances (63%), the physiological PPI weights (84%), the psychophysiological PPI weights (78%), the GC weights (53%) and the ROI activation levels during planning (75%). The best discrimination accuracy was achieved when all measures except for those from the behavioural and GC analyses were input to the SVM (NFL alumni = 84% controls = 95%).

## Discussion
Despite growing concern regarding the long-term health of sports players who suffer from repetitive mTBIs, few studies have been conducted investigating the status of retired professional players at the levels of both brain and behaviour. To our knowledge, this is the first joint computerized neuropsychological and functional magnetic resonance imaging study in a cohort of NFL alumni and the first

| Table 2 | functional connectivity analyses | | | |
|---|---|---|---|
| *General connectivity* | | | |
| **A) Collapsed** | **Con** | **t** | **P (one tailed)** |
| Left PC | 0.41 | 7.96 | **p < 0.001** |
| Right PC | 0.53 | 6.36 | **p < 0.001** |
| Left DLPFC | 0.81 | 11.95 | **p < 0.001** |
| Right FPC | 0.50 | 5.88 | **p < 0.001** |
| **B) Control > NFL alumni** | | | |
| Left PC | 0.19 | 3.66 | **p < 0.001** |
| Right PC | 0.17 | 2.02 | 0.052 |
| Left DLPFC | 0.24 | 3.56 | **p < 0.001** |
| Right FPC | 0.34 | 4.00 | **p < 0.001** |
| *Psychophysiological interaction* | | | |
| **C) Collapsed** | | | |
| Left PC | −0.13 | −1.93 | 0.063 |
| Right PC | −0.11 | −1.62 | 0.117 |
| Left DLPFC | 0.20 | 2.01 | 0.054 |
| Right FPC | 0.08 | 0.81 | 0.424 |
| **D) Control > NFL alumni** | | | |
| Left PC | 0.06 | 0.85 | 0.402 |
| Right PC | 0.06 | 0.85 | 0.403 |
| Left DLPFC | 0.05 | 0.51 | 0.613 |
| Right FPC | −0.04 | −0.42 | 0.675 |
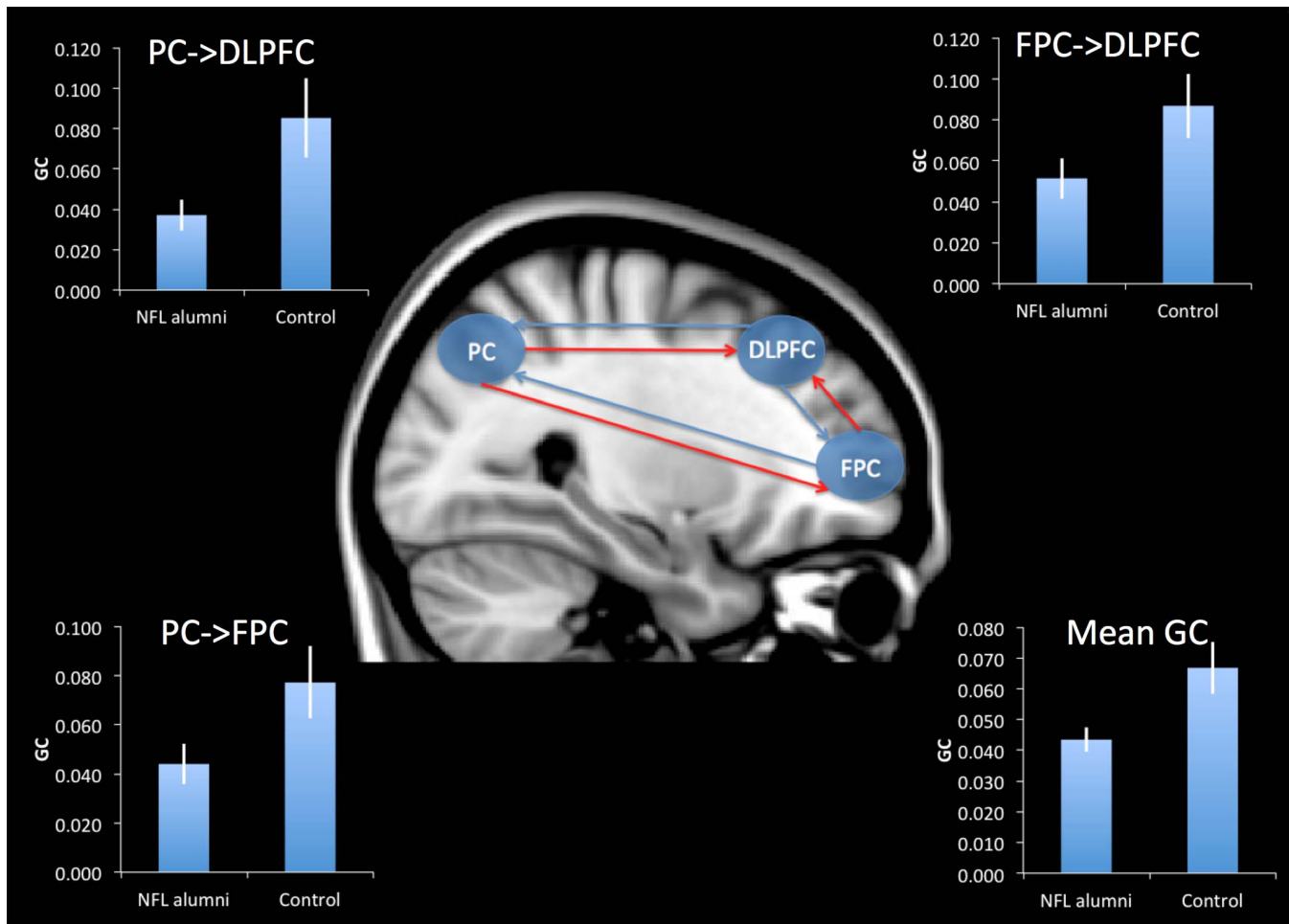
**Figure 3 | Granger Causality analysis of affective connectivity.** NFL alumni had lowered mean GC within the dorsal frontoparietal network relative to controls. This abnormal affective connectivity was characterised by a significant reduction in GC from the PC to the DLPFC, from the FPC to the DLPFC, and from the PC to the FPC. Red lines represent links with significantly reduced GC in NFL alumni relative to controls. Inset graphs display GC for NFL alumni and control groups with error bars reporting the standard error of the mean.

attempt to examine functional connectivity within this population. In accordance with previous behavioural studies, we observed only small differences for NFL alumni relative to controls when performing a computerized test of executive function. By contrast, results from the fMRI analyses revealed far more pronounced abnormalities in functional activation within the dorsal frontoparietal network. These in-vivo findings extend previous post-mortem and epidemiological studies by demonstrating that hyperfrontalility and frontoparietal hypoconnectivity are cumulative long-term consequences of repetitive mTBI.

The current study provides converging evidence to support the view that long-term cortical compensatory mechanisms work to counter the neurological impact of repetitive mTBI. For example, the NFL alumni exhibited pronounced hyperactivation within the

same DLPFC regions that were engaged by controls during planning and retrieval. This hyperactivation was evident at all levels of planning complexity and for all levels of working-memory load, including those for which there was no significant difference in performance. The profile of activation within the DLPFC was also qualitatively different between groups as it scaled significantly according to planning demands in NFL alumni but not in controls. Moreover, frontopolar regions were selectively recruited in the NFL alumni during more difficult trials whereas in controls no such activation was evident. Thus, additional cortical resources were brought on line in NFL alumni in order to cope with heavier planning demands. The observation of frontal-lobe hyperactivation is consistent with results from previous research investigating functional brain activation in college athletes during the acute phase shortly after concussion[24–27]. However, the results presented here demonstrate that across the life span, hyperfrontality either persists or resurfaces.

The functional and affective connectivity analyses provide novel insights regarding the nature of the underlying abnormality. More specifically, in the PPI analysis the NFL alumni had greatly reduced functional connectivity within the dorsal frontoparietal network. The reduction in functional connectivity was not specific to certain psychological conditions as it was consistent across planning and counting trials. The GC analysis demonstrates that hypoconnectivity was not symmetrical throughout the frontoparietal network but rather, was characterised by significant reductions in affective connectivity to the DLPFC and FPC but not to the PC. Thus, it would

| Table 3 | Granger Causality - Cross Group Analyses | | |
|---|---|---|
| | t | p (*one-tailed*) |
| **Mean GC** | 2.528 | **0.009** |
| **FPC - > DLPFC** | 2.268 | **0.017** |
| **PC - > DLPFC** | 1.907 | **0.034** |
| **DLPFC - > FPC** | 0.532 | 0.300 |
| **PC - > FPC** | 1.978 | **0.029** |
| **DLPFC - > PC** | 0.019 | 0.493 |
| **FPC - > PC** | 0.798 | 0.216 |

**5**

appear that the NFL alumni suffer from a lack of dorsal frontoparietal network coherence and abnormal causal flow. This pattern of results fits particularly closely with an inefficiency model, in which long distance connectivity within the frontoparietal network is disrupted and as a consequence, frontal lobe sub-regions work harder when dealing with executive cognitive demands.

The strength of the neuroimaging results should be contrasted with the relative weakness of the behavioural differences. More specifically, a core aim of neuropsychological testing is the development of sensitive biomarkers that can be used to confirm neurological impairments. The results of the group level analyses suggest that when quantifying the long-term neurological impact of mTBIs, functional neuroimaging can provide more sensitive biomarkers than behavioural testing alone. For example, the behavioural differences that were observed in the SP task were barely significant at the group level. This result is in line with the broader sports-mTBI literature, in which reports of longer-term impairments are typically modest and hard to replicate[21]. By contrast, the functional neuroimaging results showed robust effects in terms of both hyperactivation and hypo-connectivity in NFL alumni. The hyperactivation was evident both when examining ROIs that focused on brain regions associated with task demands and when applying a stringent voxel-wise whole brain corrected threshold. It should be noted that the ROI analyses were specifically designed to match the behavioural analyses with respect to the design and therefore, the degrees of freedom of the group-level models. Thus, the greater sensitivity of the neuroimaging analyses cannot be accounted for by any increased statistical power inherent in the factorial model. Simple scaling of the fMRI signal also cannot account for the observed results, as hyperactivation differed qualitatively across groups, and was specific to certain brain regions (i.e., DLPFC not PPC) under certain conditions (e.g. FPC under heavier planning loads). Moreover, connectivity measures were lower not higher in NFL alumni. Instead, the relative weakness of the behavioural results is most likely to be a consequence of cognitive coping strategies and cortical compensatory mechanisms that work to counter the behavioural consequences of an acquired executive impairment.

NFL alumni are a particularly difficult population to recruit for research studies of this type. However, they offer a unique opportunity to examine how individual differences in number of mTBIs suffered relate to individual differences in behavioural and neural abnormalities. Thus, it is particularly noteworthy that both hyper-frontality and hypoconnectivity were related to the total number of head collisions of great enough severity to warrant being removed from play. These correlational analyses effectively rule out lifestyle variables that may differ across the two groups as potential confounds and moreover, support the view that repetitive mTBIs have a cumulative impact in the longer term.

Indeed, more broadly, individuals who suffer from TBI often show little impairment on lab-based neuropsychological tests of executive function, yet the epidemiological data suggests that the longer-term outcomes for many such individuals are particularly poor[16]. In light of the current results, it seems likely that this inconsistency between assessments and outcomes is a consequence of similar cortical compensatory mechanisms operating in these related groups. Such compensatory mechanisms could adequately mask behavioural impairments in tests of executive function within the lab whilst failing to fully compensate in more complex real-world scenarios. In accordance with this view, it has previously been observed that the most commonly applied tests of executive function are poor predictors of real life executive impairments and therefore, lack both sensitivity and ecological validity[17,18]. This view fits particularly well with the current results, because whilst none of the NFL alumni tested here had previously been diagnosed with a cognitive impairment and cross-group behavioural differences were barely significant, many of them reported that they were experiencing distressing cognitive problems in every day life. Thus while they are not classified as patients based on behavioural assessments, we would suggest that they should be based on the conjunction of evidence from previous post-mortem studies, the self report of real life executive problems, and the observation of pronounced functional brain abnormalities within a frontoparietal network that is affected in dysexecutive neurological and psychiatric populations[19,28]. More research is required in order to determine whether similar fMRI results are evident in other populations that suffer from mTBI and to gauge their prognostic value. However, the results presented here, suggest that measures of abnormal functional activation and functional connectivity may provide a useful supplement to traditional lab based behavioural assessments when confirming impairments.

In fact, biomarkers that may be used to confirm neurological abnormalities at the individual participant level are notoriously difficult to develop but are also of great potential value. For example, a sensitive marker of frontal-lobe dysfunction might be used to confirm an acquired executive impairment when negotiating financial compensation for the long-term side effects of acquired injuries. Consequently, it is notable that the functional neuroimaging analyses were far more effective than behavioural testing when subdividing NFL alumni and controls at the individual participant level. More specifically, when the 90th percentile was taken as the criterion for abnormal performance, ~15% and ~31% of the NFL alumni were classified as low performers on the SP task using accuracy and RT as measures respectively. By contrast, ~53% were classified as functionally abnormal based on hyperactivation within the DLPFC during planning. 92% were within the bottom 10 percentile based on the functional connectivity analysis. Furthermore, whereas DLPFC functional activation and functional connectivity were found to correlate strongly with the ranked number of times players were sent off field due to head injury, behavioural data failed to show a statistically significant relationship. Thus the brain imaging data can provide more accurate detection and assessment of the level of damage acquired. In such cases, the most convincing method for confirming that a neurological abnormality has been acquired, may be the decision of a previously trained learning machine that takes into account behavioural and imaging measures together. Once trained on one cohort, the machine could be used to pass an unbiased judgement on whether an individual is most similar to a control or retired NFL player in terms of their behavioural and functional activation profiles, thereby providing a polymarker. Here, the SVM was able to make this categorical decision with between 84% and 90% accuracy depending on the set of information that was input to the machine. Critically, the distance from the hyperplane that separates controls from NFL alumni, correlated significantly with the RTO measure, supporting the view that the parameters that guided this classification decision were related to mTBI. Future longitudinal studies should determine whether this polymarker can also provide an early warning sign of accumulating damage prior to the development of behavioural symptoms in an individual in order to warn them that preventative measures should be taken.

Several important limitations should be considered in the current study. Most notably, while none of the NFL alumni had been diagnosed with or treated for neurological or psychiatric conditions, they were a self-selecting population and as outlined above, many believed themselves to be suffering from cognitive deficits acquired during the courses of their careers. The observed frontal-lobe abnormalities were almost certainly related to repetitive mTBI as demonstrated by the significant relationship with the RTO measure. However, to estimate the overall proportion of *all* NFL alumni who suffer such long-term effects would require the testing of a larger-scale randomly sampled population. This study also did not have the resolution to differentiate between concussive and sub-concussive injuries. Nor could it accurately examine the relationship between milder impacts that did not lead to removal from play and functional abnormalities

because no reliable estimate could be taken of their frequency. Determining the strength of this relationship would form a sensible focus for a longitudinal study and we suggest that the SP paradigm would provide a logical tool for this purpose.

The inherent limitations of accurately matching controls to NFL alumni should also be noted. NFL players are a unique group in many respects. By definition, they differ from controls physically and moreover, leading an NFL career will result in differences in a range of potentially confounding lifestyle factors. Indeed, we would argue that it is impossible to recruit a perfectly match set of controls for an NFL cohort. Consequently, the most appropriate way of ruling out potential confounds, is to examine individual differences in those factors that might be confounding. For example, one of the reviewers suggested that differences in baseline abilities might account for the cross-group differences in functional activation and connectivity. Measures of baseline intelligence were not available and whilst all but one participant completed college, measures of education level are hard to interpret in NFL alumni when one considers motivation and career path. It is important to note that individual differences in baseline IQ cannot be entirely ruled out in a cross-sectional study of this type. However, when individual differences in connectivity and activity were examined using both ROI and voxel-wise analyses in the control group, there was no significant relationship with task performance measures including overall accuracy and response time in ROI or voxel-wise whole brain analyses. Nor were any significant correlations observed when using education level as a predictor in the entire Applied FMRI normative cohort. Similarly, the normative controls were on average physically smaller than NFL alumni (indeed two additional NFL volunteers could not be scanned as they did not fit within the bore of the magnet). However, weight, head size, and brain dimensions did not correlate significantly with DLPFC activation in the NFL or control groups. Nor was brain size significantly different across the NFL and control groups. The Applied fMRI normative cohort also includes both males and females. However, there were no statistically significant differences between functional activation within any of the ROIs examined here or with whole brain analysis when contrasting males and females within the normative cohort. Furthermore, contrasting NFL alumni with age matched male controls only (N = 13) still generated a robust cross-group effect in terms of DLPFC hyperactivation and hypoconnectivity. In fact, of the various correlations examined, only age, which was matched across groups, showed significant effects in controls, with a modest correlation in terms of DLPFC-PC connectivity (left $t = 1.77$ $p < 0.05$, right $t = 1.72$ $p = 0.05$). These modest effects of age, were dwarfed by the scale of the NFL-control effects and more importantly, stand in stark contrast to the reliability of the correlations with the RTO measure. Thus, whilst confounding variables cannot be entirely ruled out in a cross sectional study of this type, when the individual differences and cross group analyses are considered together, the results accord particularly closely with the hypothesis that exposure to repetitive mTBI during the course of the professional NFL career leads to abnormal frontal lobe function.

In summary, the results presented here provide novel evidence that cortical compensatory mechanism work to counteract disrupted dorsal executive network connectivity in purportedly healthy NFL alumni who have suffered repetitive mTBI during the course of their careers. These changes are important in light of the inconsistencies in the neuropsychological literature and the increased incidence of neurological disease and neurodegenerative mortality among NFL alumni and sufferers of mild to moderate TBI more broadly. The relationship between history of mTBI, abnormal frontoparietal function, and long-term outcomes clearly merits further investigation. To this end, we believe that larger scale and longitudinal neuroimaging studies are warranted in this and other populations that suffer from multiple mTBIs. Such studies have the potential to derive the neural fingerprint of mTBI in order to provide an early warning sign when repeated injuries are accumulating towards long-term cognitive impairment.

## Methods

This study was approved by the Pomerado Hospital IRB in San Diego, CA, USA. All participants gave informed consent prior to taking part in the study.

**Participants.** The Applied fMRI normative cohort consists of 60 healthy volunteers recruited from the San Diego area. All volunteers are right-handed college graduates, fluent in English, with normal or corrected-to-normal eyesight, normal hearing, and no history of neurological or psychiatric illnesses. These individuals were selected to cover a range of ages, between 20 and 60. In the current study, an age-matched subset of 20 volunteers was compared to the NFL group (mean age controls = 53, NFL alumni = 54). Thirteen retired American professional football players participated in the neuroimaging study. They had no diagnosis of neurological or psychiatric illness. A further 2 retired players volunteered but could not be included in the neuroimaging study because they did not fit within the bore of the scanner. No clinically significant MRI structural abnormalities were identified in any of the participants. All participants were right handed, had normal or corrected-to-normal eyesight and normal hearing. NFL alumni were asked to report the total number of times that they had been taken out of play due to head impact, with this measure including instances in which they returned to the game and instances in which they remained out of play. Baseline measures of IQ and education achievement were not available for the NFL cohort.

**Executive paradigm.** The fMRI task has been reported in detail elsewhere[19]. In brief, on a given trial, the participant was presented with two arrays of coloured balls held in transparent vertical tubes (**Fig. 1a**). During planning trials, the participant was required to work out the minimum number of moves that would be required to rearrange the balls in the array at the top of the screen so that they were in the same configuration as those in the array at the bottom of the screen. They were instructed that they could only move the balls in and out of the top of the tubes and could only move one ball at a time. Once they had calculated the minimum number of moves required, they reported their answer by pressing the corresponding button on an fMRI compatible button box that was placed beneath the right hand. The difficulty of the planning trials was varied from 2 to 4 moves. There were also control trials, in which the same arrays were displayed but with varying numbers of balls in the upper and lower panels. During the control trials, participants simply subtracted the number of balls in the top array from the number in the bottom and indicated the resultant value using the same button box response. This condition was intended to control for visual, attentional and motor demands and was used in a subtractive analysis to identify brain regions that were particularly active during planning. Trials were pseudo-randomised according to a predefined sequence. Stimulus arrays were displayed until the participant responded, subsequent to which there was a 10 second period of rest. The task ran for a total of 10 minutes and participants were required to accurately solve as many problems as possible within this time. Prior to entering the scanner, participants undertook a pre-training session, which involved reading the task instructions and undertaking a short series of practice trials. The fMRI data were modelled at the individual participant level using fixed effects analysis in SPM8. Six predictor functions were generated, corresponding to the onsets and duration of planning and counting trials convolved with the canonical haemodynamic response function separately for trials for which the answer was 2, 3 or 4. Noise due to movement was accounted for by inclusion of six nuisance variables corresponding to translations and rotations in the x, y, and z planes. The whole brain maps, depicting beta weights for the six experimental predictor functions, were collated for group level whole brain and regions of interest (ROI) analyses.

**Neuroimaging.** FMRI data were collected using the standard 2 second EPI protocol on a Siemens 3T Tim Trio scanner at the Applied fMRI Institute in San Diego, USA. The data were pre-processed and analysed in SPM8 (Statistical Parametric Mapping, Welcome Department of Imaging Neuroscience, London, UK). Prior to analysis, data were slice-timing and motion corrected, spatially normalized to the standard Montreal Neurological Institute template and spatially smoothed with an 8 mm full width at half maximum Gaussian kernel. The data were high-passed filtered (cutoff period = 180 seconds) to remove low frequency drifts in the MRI signal.

In order to examine task related functional activation, fixed effects analyses were carried out at the individual-subject level using general linear modelling. Focused regions of interest (ROIs) analyses were carried out with the MarsBaR toolbox[29], which calculates the average intensity value from all voxels within a ROI. Voxel-wise group level analyses were carried out in SPM8 and unless reported otherwise used a False Discovery Rate (FDR) correction of $P < 0.05$ for multiple comparisons across the whole brain volume. Ranked Times Off (RTO) was calculated when examining individual differences because the number of times that the NFL alumni reported having been taken out of play after head impact was positively skewed. Non-parametric functional imaging analyses used Ranked Times Off (RTO) as a predictor function and were conducted using the Free Surf Library randomise function with voxel-wise thresholding at $p < 0.01$. False positives were controlled at the cluster level[30,31] using family wise error (FWE) correction at $p < 0.05$ within the task localiser mask, which included all regions that were more active for task relative to rest collapsed across group at $p < 0.001$ uncorrected.

Functional connectivity was examined using psychophysiological interaction analyses in SPM8. For each participant, timecourse data were extracted from a 10 mm radius spherical ROI using the Volume of Interest function, which extracts the first eigenvector from all voxels within the ROI. The extracted timecourse was deconvolved in order to estimate underlying neural activation from the BOLD response. The estimated timecourse of neural activation was interacted with a psychological timecourse for planning vs. counting trials. The resultant psychological, physiological and psychophysiological timecourses were convolved with the haemodynamic response function to generate a predictor of BOLD activation and these predictors were entered into a general linear model in SPM8 along with six movement variables corresponding to translations and rotations in the x, y, and z planes. The beta maps for the physiological and psychophysiological predictor functions were examined at the group level using independent sample t-tests in MarsBaR.

Affectivity connectivity was examined by Granger Causality analyses with the Granger Causality Connectivity Analysis (GCCA) toolbox[22,23]. Mean timecourses were extracted from ROIs using MarsBaR. The timecourse were detrended and demeaned in GCCA. Optimal model orders (the number of previous observations to include in the GC model) were calculated separately for each individual using Bayesian information criteria (BIC) and the mean order calculated across participants. Conditional G-Causality models were fitted for each individual with lagged timecourses set to the mean BIC of 3 to ensure that model order was evenly matched across all individuals. The resultant GC magnitudes for each of the six connections were extracted for cross group analysis and GC magnitudes above 2.5 SDs from the mean replaced with the next nearest non-outlier value for the same connection in the same group. Mean GC was calculated by averaging the magnitudes for all six connections. Cross-group comparisons of GC magnitudes were conducted using independent samples t-tests with variance not assumed to be equal in SPSS.

SVM classification was undertaken using the svmtrain and svmclassify functions, which are native to Matlab. Validation used a leave-one-out approach whereby one participant data set was excluded from each round of training subsequent to which the trained SVM attempted to classify the excluded data set, to which it was naïve. This process was repeated until each individual data set had been left out and classified once in order to generate an unbiased measure of classification accuracy. The SVM used all of Matlab's default parameters including the linear kernel function (dot product) and the sequential minimal optimisation method.

1. Russo, M. B. & Crutchfield, K. E. Ethical perspectives in neurology: sports-acquired traumatic brain injury. *Continuum (Minneap Minn)* **16**, 161–165, doi:10.1212/01.CON.0000391462.98911.26 (2010).
2. Langlois, J. A., Rutland-Brown, W. & Wald, M. M. The epidemiology and impact of traumatic brain injury: a brief overview. *J Head Trauma Rehabil* **21**, 375–378 (2006).
3. Crisco, J. J. et al. Frequency and location of head impact exposures in individual collegiate football players. *J Athl Train* **45**, 549–559, doi:10.4085/1062-6050-45.6.549 (2010).
4. Crisco, J. J. et al. Head impact exposure in collegiate football players. *J Biomech* **44**, 2673–2678, doi:10.1016/j.jbiomech.2011.08.003 (2011).
5. Greenwald, R. M., Gwin, J. T., Chu, J. J. & Crisco, J. J. Head impact severity measures for evaluating mild traumatic brain injury risk exposure. *Neurosurgery* **62**, 789–798; discussion 798, doi:10.1227/01.neu.0000318162.67472.ad (2008).
6. Marchi, N. et al. Consequences of repeated blood-brain barrier disruption in football players. *PLoS One* **8**, e56805, doi:10.1371/journal.pone.0056805 (2013).
7. Pellman, E. J., Lovell, M. R., Viano, D. C. & Casson, I. R. Concussion in professional football: recovery of NFL and high school athletes assessed by computerized neuropsychological testing–Part 12. *Neurosurgery* **58**, 263–274; discussion 263–274, doi:10.1227/01.NEU.0000200272.56192.62 (2006).
8. Pellman, E. J., Lovell, M. R., Viano, D. C., Casson, I. R. & Tucker, A. M. Concussion in professional football: neuropsychological testing--part 6. *Neurosurgery* **55**, 1290–1303; discussion 1303–1295 (2004).
9. Guskiewicz, K. M. et al. Association between recurrent concussion and late-life cognitive impairment in retired professional football players. *Neurosurgery* **57**, 719–726; discussion 719–726 (2005).
10. Lehman, E. J., Hein, M. J., Baron, S. L. & Gersic, C. M. Neurodegenerative causes of death among retired National Football League players. *Neurology* **79**, 1970–1974, doi:10.1212/WNL.0b013e31826daf50 (2012).
11. Weir, D. R., Jackson, J. S. & Sonnega, A. National Football League Player Care Foundation: Study of retired NFL players. (Institute for Social Research University of Michigan, Ann Arbor, MI, 2009).
12. Omalu, B. I. et al. Chronic traumatic encephalopathy in a National Football League player. *Neurosurgery* **57**, 128–134; discussion 128–134 (2005).
13. McKee, A. C. et al. Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury. *J Neuropathol Exp Neurol* **68**, 709–735, doi:10.1097/NEN.0b013e3181a9d503 (2009).
14. McKee, A. C. et al. The spectrum of disease in chronic traumatic encephalopathy. *Brain* **136**, 43–64, doi:10.1093/brain/aws307 (2013).
15. Amen, D. G. et al. Impact of playing American professional football on long-term brain function. *J Neuropsychiatry Clin Neurosci* **23**, 98–106, doi:10.1176/appi.neuropsych.23.1.98 (2011).
16. McMillan, T. M., Teasdale, G. M., Weir, C. J. & Stewart, E. Death after head injury: the 13 year outcome of a case control study. *Journal of neurology, neurosurgery, and psychiatry* **82**, 931–935, doi:10.1136/jnnp.2010.222232 (2011).
17. Hanna-Pladdy, B. Dysexecutive syndromes in neurologic disease. *J Neurol Phys Ther* **31**, 119–127, doi:10.1097/NPT.0b013e31814a63c2 (2007).
18. Wood, R. L. & Liossi, C. The ecological validity of executive tests in a severely brain injured sample. *Arch Clin Neuropsychol* **21**, 429–437, doi:10.1016/j.acn.2005.06.014 (2006).
19. Williams-Gray, C. H., Hampshire, A., Robbins, T. W., Owen, A. M. & Barker, R. A. Catechol O-methyltransferase Val158Met genotype influences frontoparietal activity during planning in patients with Parkinson's disease. *J Neurosci* **27**, 4832–4838 (2007).
20. Shallice, T. Specific impairments of planning. *Philos Trans R Soc Lond B Biol Sci* **298**, 199–209 (1982).
21. Belanger, H. G., Spiegel, E. & Vanderploeg, R. D. Neuropsychological performance following a history of multiple self-reported concussions: a meta-analysis. *J Int Neuropsychol Soc* **16**, 262–267, doi:10.1017/S1355617709991287 (2010).
22. Bressler, S. L. & Seth, A. K. Wiener-Granger causality: a well established methodology. *Neuroimage* **58**, 323–329, doi:10.1016/j.neuroimage.2010.02.059 (2011).
23. Seth, A. K. A MATLAB toolbox for Granger causal connectivity analysis. *J Neurosci Methods* **186**, 262–273, doi:10.1016/j.jneumeth.2009.11.020 (2010).
24. Chen, J. K. et al. Functional abnormalities in symptomatic concussed athletes: an fMRI study. *Neuroimage* **22**, 68–82, doi:10.1016/j.neuroimage.2003.12.032 (2004).
25. Jantzen, K. J., Anderson, B., Steinberg, F. L. & Kelso, J. A. A prospective functional MR imaging study of mild traumatic brain injury in college football players. *AJNR Am J Neuroradiol* **25**, 738–745 (2004).
26. Pardini, J. E. et al. Postconcussive symptoms are associated with compensatory cortical recruitment during a working memory task. *Neurosurgery* **67**, 1020–1027; discussion 1027–1028, doi:10.1227/NEU.0b013e3181ee33e2 (2010).
27. Slobounov, S. M. et al. Functional abnormalities in normally appearing athletes following mild traumatic brain injury: a functional MRI study. *Exp Brain Res* **202**, 341–354, doi:10.1007/s00221-009-2141-6 (2010).
28. Grant, J. E. et al. A Proof of Concept Study of Tolcapone for Pathological Gambling: Relationships with COMT Genotype and Brain Activation. *European Neuropsychopharmacology* (2013).
29. Brett, M., Anton, J., Valabregue, R. & Poline, J. in *8th International Conference on Functional Mapping of the Human Brain*.
30. Bullmore, E. T. et al. Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Trans Med Imaging* **18**, 32–42 (1999).
31. Hayasaka, S. & Nichols, T. E. Validating cluster size inference: random field and permutation methods. *Neuroimage* **20**, 2343–2356 (2003).
32. Hampshire, A., Highfield, R. R., Parkin, B. L. & Owen, A. M. Fractionating human intelligence. *Neuron* **76**, 1225–1237, doi:10.1016/j.neuron.2012.06.022 (2012).

## Acknowledgements

## Author contributions

A.H. conceived the project, designed and programmed the fMRI paradigm, designed and conducted the analysis approach, and produced the final draft of the manuscript. A.M. contributed to the fMRI analyses, conducted the literature search and produced the first draft of the manuscript. A.M.O. conceived the project, designed the fMRI paradigm and edited the final draft of the manuscript.

## Additional information