NeuroImage: Clinical 23 (2019) 101886

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Facial emotion recognition in children treated for posterior fossa tumours and typically developing children: A divergence of predictors

Iska Moxon-Emre[a,b,c], Norman A.S. Farb[b], Adeoye A. Oyefiade[a,b], Eric Bouffet[a,b], Suzanne Laughlin[a,b], Jovanka Skocic[a], Cynthia B. de Medeiros[a], Donald J. Mabbott[a,b,*]

[a] The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada
[b] University of Toronto, Toronto, ON M5S 3G3, Canada
[c] Pediatric Oncology Group of Ontario, Toronto, ON M5G 1V2, Canada

ARTICLE INFO

Keywords:
Facial emotion recognition
Brain tumours
Diffusion tensor imaging
White matter
Eye-tracking
Development

ABSTRACT

Facial emotion recognition (FER) deficits are evident and pervasive across neurodevelopmental, psychiatric, and acquired brain disorders in children, including children treated for brain tumours. Such deficits are thought to perpetuate challenges with social relationships and decrease quality of life. The present study combined eye-tracking, neuroimaging and cognitive assessments to evaluate if visual attention, brain structure, and general cognitive function contribute to FER in children treated for posterior fossa (PF) tumours (patients: $n = 36$) and typically developing children (controls: $n = 18$). To assess FER, all participants completed the Diagnostic Analysis of Nonverbal Accuracy (DANVA2), a computerized task that measures FER using photographs, while their eye-movements were recorded. Patients made more FER errors than controls ($p < .01$). Although we detected subtle deficits in visual attention and general cognitive function in patients, we found no associations with FER. Compared to controls, patients had evidence of white matter (WM) damage, (i.e., lower fractional anisotropy [FA] and higher radial diffusivity [RD]), in multiple regions throughout the brain (all $p < .05$), but not in specific WM tracts associated with FER. Despite the distributed WM differences between groups, WM predicted FER in controls only. In patients, factors associated with their disease and treatment predicted FER. Our study provides insight into predictors of FER that may be unique to children treated for PF tumours, and highlights a divergence in associations between brain structure and behavioural outcomes in clinical and typically developing populations; a concept that may be broadly applicable to other neurodevelopmental and clinical populations that experience FER deficits.

1. Introduction

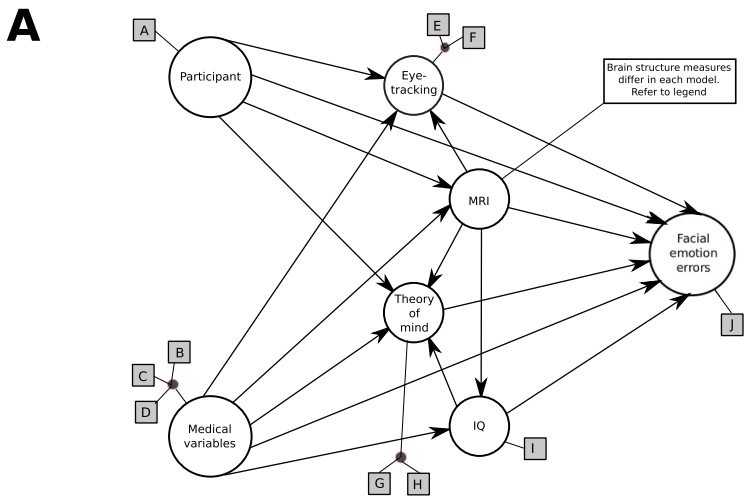
Facial emotions provide rich non-verbal information in real-time, and their correct interpretation is critical to participation in one's social environment (Collin et al., 2013). Facial emotion recognition (FER) deficits are evident and pervasive across various disorders in children; namely, psychiatric disorders (e.g., schizophrenia/psychosis, mood disorders and anxiety [reviewed in Collin et al., 2013]), neurodevelopmental disorders (e.g., autism [Eussen et al., 2015; Evers et al., 2015; Taylor et al., 2015] and ADHD - although the evidence is mixed [Bunford et al., 2015]), neurological conditions (e.g., epilepsy [Edwards et al., 2017]), and acquired brain disorders (e.g., traumatic brain injury [Mancuso et al., 2015] and brain tumours [Bonner et al., 2008]). FER deficits have been proposed to perpetuate the challenges with social relationships that individuals with many of these disorders experience (Collin et al., 2013). Facial emotions provide bi-directional information between individuals, and FER errors can result in miscommunications about intentions and reactions; thus, the ability to recognize facial emotions is thought to be required for effective social participation (Morris et al., 2009). Factors that contribute to FER remain poorly understood in children, yet their identification is necessary in the effort to mitigate the deleterious effects of FER deficits.

Here, we consider three factors that may contribute to FER: visual attention, brain structure and general cognitive function. Children treated for tumours that arise in the posterior fossa (PF) of the brain may provide insight on how FER is disrupted; this clinical population has a documented FER deficit (Bonner et al., 2008) and they experience: 1) attention problems, 2) altered brain structure (i.e., white matter [WM] and cerebellar damage), and 3) cognitive deficits (reviewed in Padovani et al., 2012).

* Corresponding author at: The Hospital for Sick Children, 686 Bay Street, Toronto, ON M5G 0A4, Canada.
E-mail address: donald.mabbott@sickkids.ca (D.J. Mabbott).

https://doi.org/10.1016/j.nicl.2019.101886
Received 19 October 2018; Received in revised form 9 May 2019; Accepted 28 May 2019
Available online 12 June 2019

Full Model Structure - all paths tested

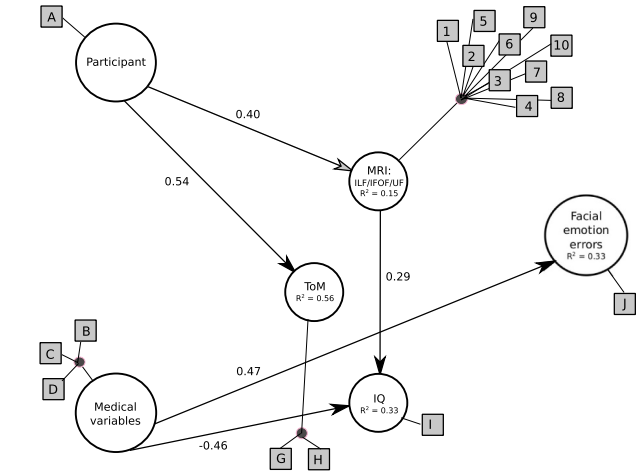


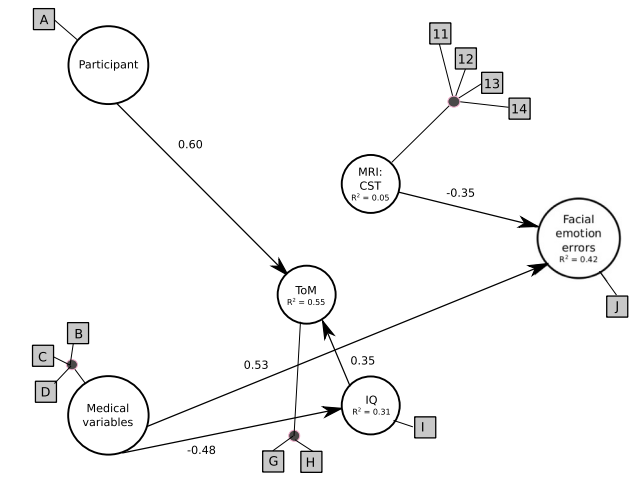
Latent Construct Indicators:

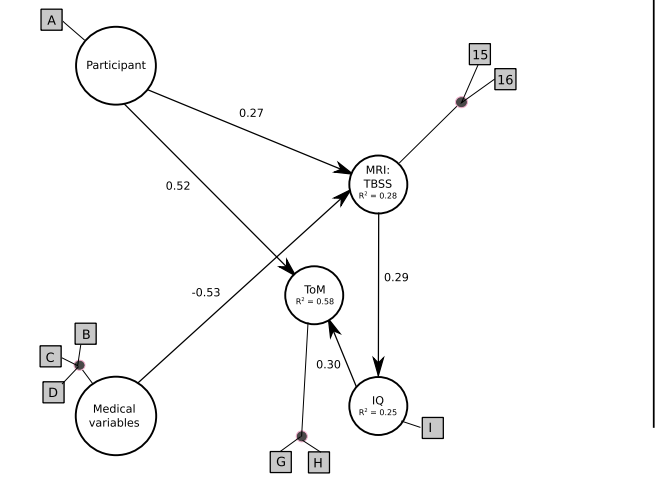
A = Age at testing
B = NPS score
C = Complications score
D = Time since diagnosis
E = Number of fixations
F = Total dwell time
G = Advanced ToM
H = EEFT total score
I = IQ
J= Incorrect responses (DANVA2)

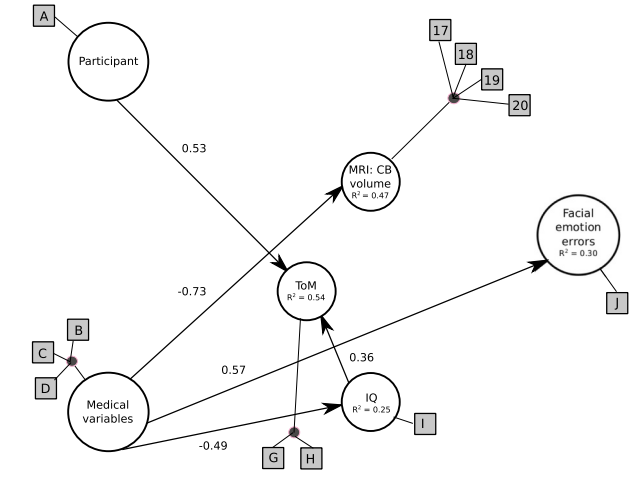
1-20: MRI metrics used. Differs in each model. Refer to legend

Model Results - significant paths only









(caption on next page)

Fig. 1. Our PLS path models designed to test if visual attention, brain structure and general cognitive function contribute to FER. A, The full model structure, showing all paths tested. We predicted that age and medical variables will have direct and indirect effects on the number of FER errors, through visual attention (i.e., eye-tracking), brain structure (differs according to model tested) and general cognitive function (i.e., theory of mind and/or IQ) processes. B-E, four PLS path models tested with only the significant paths and standardized path coefficients shown. Indicators that contribute to the latent constructs remain unchanged in all models, except for those that contribute to the brain structure latent construct, as follows: *B, model 1:* 1–4 = FA and RD of the left and right IFOF; 5–8 = FA and RD of the left and right ILF; 9–10 = RD of the left and right UF. *C, model 2:* 11–14: FA and RD of the left and right CST. *D, model 2:* 15–16: FA ad RD from voxels that differed significantly between patients and healthy controls in the 2-group TBSS analysis. *E, model 4:*17–20: cerebellar grey and WM volume in left and right hemisphere. NPS = neurological predictor scale; ToM = theory of mind; EEFT = emotional and emotive faces task; IQ = intelligence quotient; DANVA2 = diagnostic analysis of nonverbal accuracy.

Children treated for brain tumours may experience FER deficits for a number of reasons. First, they may not effectively attend to visual stimuli such as the information present in faces, including emotion. To evaluate visual attention, we recorded participant's eye-movements while they performed a FER task. Visual information is perceived primarily by fixating on regions of interest; multiple times every second, a viewer selects, via a saccadic movement, a region for perceptual and cognitive processing (Henderson and Hollingworth, 1998). Measuring eye-movements has been shown to be reliable index of visual attention (Duc et al., 2008). Second, alterations to their brain structure may contribute to FER deficits. Studies in adults with brain lesions, traumatic brain injury (reviewed in Wang et al., 2018) and healthy adults (Coad et al., 2017) suggest the following WM tracts are associated with FER: the inferior-frontal occipital fasciculus (IFOF) and the inferior-longitudinal fasciculus (ILF), tracts that connect the occipital cortex to the orbitofrontal cortex and anterior temporal lobe respectively, and the uncinate fasciculus (UF), tracts that connect the anterior temporal lobe to the orbitofrontal cortex (Catani and Thiebaut de Schotten, 2012). Analogous studies are lacking in children; thus, we associated FER abilities with microstructure of ILF, IFOF and UF. We also took an unbiased approach to consider the relations between FER and WM throughout the entire brain. We used diffusion tensor imaging (DTI), a technique that generates quantifiable indices based on the directionality and displacement of water that are thought to reflect WM organization (Basser, 1995; Jones and Leemans, 2011). Fractional anisotropy (FA) reflects the direction of principal diffusion, and is though to reflect functionally relevant properties of WM microstructure (Beaulieu, 2002). Radial diffusivity (RD), also referred to as perpendicular diffusion, is thought to reflect myelin architecture (Song et al., 2002, 2005). The cerebellum has also been implicated in emotion recognition (reviewed in Clausi et al., 2017), and it is important for oculomotor control (Beh et al., 2017); thus we evaluated grey and white matter volumes in the cerebellum. Lastly, general cognitive problems in children treated for PF tumours may be a contributing factor to FER deficits. General cognitive ability correlates positively with FER in typically developing children (Lawrence et al., 2015) as does theory of mind (ToM); the ability to perspective-take and infer mental states and behaviours of others (Lee et al., 2014; Cotter et al., 2018). To evaluate general cognitive function, we administered an abbreviated intelligence test (Wechsler, 2011) and two ToM measures (Hutchins et al., 2011; Dennis et al., 2013). Given that FER deficits can contribute to social problems, we also evaluated parent-reported social functioning in our sample.

In this study, we evaluated if visual attention, brain structure and general cognitive function contribute to FER in typically developing children and children treated for PF tumours, aged 8–17 years. Throughout this age range, eye-movement control improves (Luna et al., 2008), WM is actively maturing (Lebel and Beaulieu, 2011), and improvements in ToM (Im-Bolter et al., 2016) and FER (Herba and Phillips, 2004) abilities are evident. Given that these measures have developmental trajectories, it was important to control for age where appropriate. To our knowledge, microstructure of the ILF, IFOF and UF have not yet been associated with FER abilities in children. It is important to consider that established associations between brain structure and behavioural outcomes in adults may not be directly applicable to children or to pediatric clinical populations. However, given that the uninjured brain is on a developmental trajectory towards that of the adult brain, it is possible that WM tracts in children subserve the same function as adults. In contrast, children with brain injury may experience an altered developmental trajectory, and unique associations between brain structure and behavioural outcomes may result. It is thus critical to evaluate the associations between FER abilities and WM microstructure of pathways that have been implicated in FER in adults, both in typically developing children, and in children with brain injury.

We began by evaluating how visual attention, brain structure and general cognitive function differed in our typically developing and clinical sample. Next, we tested variations of a model to evaluate the relations between these factors, and to examine if they predicted FER. We expected that: 1) the WM tracts associated with FER in adults would be implicated in FER abilities in all children, 2) medical variables associated with treatment for PF tumours would have direct and indirect effects on FER, via WM tracts known to be involved in FER, and potentially also through its effects on visual attention and general cognitive function. Given that visual attention and general cognitive function may be influenced by WM microstructure, these factors were also considered as mediators between WM and FER. The full model structure, and all paths tested, are provided in Fig. 1A (top panel). Next, we tested variations on this model as secondary exploratory analyses; this was done to evaluate if WM damage throughout the brain (i.e., on a voxelwise basis) or damage to the cerebellum (i.e., cerebellar grey and WM volumes), instead of the abovementioned WM tracts, contribute to FER. If certain factors predict FER deficits in our clinical sample, we may acquire unique insights into the pathology associated with FER deficits experienced by children with various developmental and clinical disorders alike.

2. Materials and methods

2.1. Participants

Fifty-four youth between the ages of 8 and 17 completed this study; 36 children treated for PF tumours (17 patients treated with surgery with or without chemotherapy, and 19 patients treated with surgery, chemotherapy and radiation) at the Hospital for Sick Children (SickKids; Toronto, Canada) and 18 healthy control children. 6 participants did not undergo MRI, but completed all other components of the study. Participant demographic and medical variables are summarized in Table 1. Five controls were siblings of PF tumour patients who completed the present study, and another five controls were siblings of brain tumour patients who participated in other studies at the hospital (i.e., non-PF tumour patients). Thus, 10/18 (56%) of our control sample was related to brain tumour patients. All patients were > 1 year post diagnosis and had completed all therapy. Patients were excluded from participation if they had premorbid neurological disorders, or if they were receiving palliative care. Healthy controls were (self-described as) free of all neurological or clinical disorders. Patients were informed of this study via mailed letters, and/or were approached during their routine clinic visits when applicable. Healthy controls were either recruited from the community, were siblings of patients, or family members of SickKids staff. This study was approved by the hospital's Research Ethics Board. Prior to participation, parents provided written informed consent and children provided assent. When deemed capable

Table 1
Participant characteristics and patient treatment information.

	Healthy control	Surgery	Radiation	*p* value
	n = 18	*n* = 17	*n* = 19	
Sex (male)	6	12	11	0.08
Average parental education (years)				0.10
Mean	17.72	16.66	15.67	
Standard deviation	3.06	2.56	2.18	
Range	12.5–23.5	13.0–23.0	12.0–20.0	
Age at assessment (years)				0.09
Mean	12.29	12.74	14.13	
Standard deviation	2.44	3.13	2.21	
Range	8.1–16.7	8.3–17.9	8.3–16.7	
MRI scan (3 T)	17	14	17	0.52
Age at diagnosis (years)	–			0.26
Mean	–	6.23	7.43	
Standard deviation	–	3.65	2.59	
Range	–	1.8–15.4	3.0–12.2	
Time since diagnosis (years)	–			0.97
Mean	–	6.66	6.70	
Standard deviation	–	2.97	2.97	
Range	–	1.4–11.3	1.3–11.4	
Tumour size (mm^2)[a]	–			0.54
Mean	–	2124.29	1883.44	
Standard deviation	–	1208.86	916.86	
Range	–	506–4800	624–3996	
Tumour type	–			< 0.001
Medulloblastoma (average risk)	–	1	11	
Medulloblastoma (high risk)	–	1	4	
Ependymoma	–	1	4	
Pilocytic astrocytoma	–	13	0	
Cribriform neuroepithelial Tumour	–	1	0	
Tumour location within the posterior fossa	–			0.08
Midline	–	9	16	
Left hemispheric	–	2	0	
Right hemispheric	–	5	2	
Unavailable	–	1	1	
Gross total resection (> 95% of tumour resected)	–	13	16	0.43
Hydrocephalus	–			
No Hydrocephalus	–	6	1	0.03
Hydrocephalus with no treatment (resolved)	–	5	5	0.56
Hydrocephalus requiring CSF diversion (EVD, shunt, ventriculostomy)	–	6	12	0.09
Mutism following surgery[b]	–	3	7	0.18
Neurological complications	–			
Cranial nerve deficit	–	1	5	0.12
Hearing Loss	–	2	8	0.37
Meningitis	–	0	2	0.27
Motor deficits (ataxia, dysmetria, dysdiadochkinesia)	–	6	19	0.02
Visual impairment (Nystagmus; diplopia)	–	7	10	0.36
Recurrence	–			0.45
0		15	14	
1		2	4	
2		0	1	
Number of surgeries	–			0.24
1	–	16	14	
2	–	1	4	
3+	–	0	1	
Radiation type	–			< 0.001
None	–	17	0	
Focal (5400–5940 cGy)	–	0	3	
Reduced dose Cranial-Spinal (2340 cGy) + TB Boost (3240 cGy)	–	0	11	
Standard dose Cranial-Spinal (2340–3600 cGy) + TB or PF Boost (1800–3240 cGy)	–	0	5	
Chemotherapy	–			0.01
None	–	12	5	
ACNS0332 (carboplatin, cyclophosphamide, vincristine, cisplatin, isotretinoin)	–	0	1	
COG9961 (vincristine, lomustine, cisplatin)	–	0	2	
COG99703 (thiotepa, carboplatin)		2	1	
Head Start II (vincristine, cisplatin, cyclophosphamide, etoposide, methotrexate)	–	1	0	
POG9631 (etoposide, cisplatin, cyclophosphamide, vincristine)	–	0	1	
SJMB96 & SJMB03 (vincristine, cisplatin, cyclophosphamide)	–	0	9	
Vinblastine monotherapy	–	2	0	

[a] Tumor size was unavailable for 6 patients (3 surgery, 3 radiation).
[b] Patients were classified as having mutism if they had diminished speech output, linguistic difficulties or dysarthria following surgery. Mutism is a transient dysfunction and had resolved in all participants by the time of baseline assessment.

to do so, participants (typically adolescents) provided their own written consent.

2.2. Patient treatment information

All patients had surgical intervention for their PF tumours, and a subset of patients received chemotherapy with or without radiation. We divided our patient sample into two groups - patients who received radiation and patients who did not (herein referred to as radiation and surgery groups, respectively). In the radiation group, patients treated with photon beam CSI received either standard (3060–3940 centigray (cGy)) or reduced (1800–2340 cGy) dose, and a boost to the tumour bed, whereas patients treated with focal radiation received 5400–5940 cGy to their tumour site. Medical variables for these two treatment groups are summarized in Table 1. The patient groups did not differ on most medical variables, except that the radiation group had fewer patients without hydrocephalus ($p = .03$, Fisher's exact test), and more patients with motor deficits ($p = .02$, Fisher's exact test). The only other differences between the groups arose from factors that determined their groupings; namely, patients in the radiation group were diagnosed with metastatic PF tumours (either medulloblastoma or ependymoma), whereas patients in the surgery group were primarily diagnosed with benign PF tumours (most commonly pilocytic astrocytoma), or patients were younger than 3 years at diagnosis and consequently did not receive radiation ($p < .0001$, Fisher's exact test). Moreover, the chemotherapy protocols ($p = .01$, Fisher's exact test) and the radiation type differed between groups ($p < .0001$, Fisher's exact test).

2.3. FER: diagnostic analysis of nonverbal accuracy (DANVA2) task

To assess FER, all participants completed the DANVA2, a computerized task using 48 photographs of males and females (24 children and 24 adults) (Baum and Nowicki, 1998). Participants were asked to look at each photograph and to decide if the individual felt happy, sad, angry, or fearful (scared). Each photograph was displayed on the screen, along with four response boxes indicating the emotions to choose from. The photograph remained on the screen for 2000 ms, whereas the response boxes remained on the screen until a decision about the emotion was made. In order for eye-movements to be recorded during this task (to assess visual attention), an eye-tracking version of the DANVA2 task was created, using the raw images from the original computer program (Baum and Nowicki, 1998).

2.4. Visual attention

Eye-tracking apparatus and setup for DANVA2 task.

Eye movements were recorded throughout the entire DANVA2 task using a SR Research Ltd. Eyelink 1000 plus (Mississauga, Canada) eye-tracking desktop monocular system. The right eye was tracked in all except 9% ($n = 5$) participants, where poor calibration with the right eye prompted a switch to tracking the left eye instead. A sampling rate of 500 Hz and a spatial resolution of 0.01° was used. A 9-point calibration was performed prior to the experiment, and was successfully achieved for all participants. Images were displayed on a 14.5 × 12.5 in. LCD monitor with a 1280 × 1024 pixel resolution. Photographs displaying facial emotions were 1425 × 810 pixels in size, and response boxes listing the emotions were 173 × 53 pixels in size. Given that damage to the cerebellum can cause visual impairments, interest areas for eye-tracking analysis were deliberately kept large by placing them around the entire photograph; this approach was taken to avoid small alterations in eye position spatial resolution from obscuring our results. Participants were seated 26 in. from the monitor and a chin rest was used to limit head movements. The experiment was built using the Experiment Builder software provided with the SR Research Ltd. eye-tracker. Measures of visual attention included the number of fixations, and total time spent looking at the photograph (i.e., total dwell time) during the FER task; eye-tracking measures that have been utilized to evaluate gaze patterns when viewing emotional images (Calvo and Lang, 2004).

2.5. Brain structure

2.5.1. Neuroimaging protocol

Magnetic Resonance Imaging (MRI) was performed at SickKids using a Siemens 3 T whole-body MRI scanner (Prisma fit) with a 12-channel head coil. Imaging included a T1 AX 3D MPRAGE Grappa 2 protocol (T1 = 900 ms, TE/TR = 3.83/2300 ms, 160 contiguous axial slices, flip angle = 9°, 256 × 224 matrix, FOV = 256x224mm, voxel size = 1 mm ISO) and diffusion-weighted single shot spin echo DTI sequence with EPI readout (30 directions, b = 1000s/mm^2, TE/TR = 90/9000 ms, 70 contiguous axial slices, flip angle = 90°, 122 × 122 matrix interpolated to 244 × 244, FOV = 244x244mm, voxel size = 2 mm ISO, interpolated to 1x1x2mm).

2.5.2. MRI pre-processing

Cortical reconstruction and volumetric segmentation of the anatomical T1 images was performed with the FreeSurfer image analysis suite, as documented online (http://surfer.nmr.mgh.harvard.edu/), and we parcellated the cortex into 164 brain regions using a well-validated cortical atlas (Fischl et al., 2004; Destrieux et al., 2010). Although the utility of this parcellation was primarily for the additional processing steps detailed below, cerebellar grey and WM volumes from this parcellation were included in our analysis. Diffusion weighted images (DWI) were denoised, eddy current corrected for current distortions, motion corrected and bias corrected to correct B1 field inhomogeneities, using MRTrix3 package (www.mrtrix.org). All DWI images were non-linearly registered to T1 (Talairach) space, using ANTs (Avants et al., 2011). The automated cortical reconstruction and volumetric segmentation steps were not successful in 3 patients, likely as a result of their atypical anatomy (enlarged ventricles and lack of an intact cerebellum); they were not included in the probabilistic tractography analyses as a result.

2.5.3. Probabilistic tractography

Fibre orientation distributions (FOD) were estimated using a constrained spherical deconvolution (CSD) model (Tournier et al., 2013), DTI index maps (FA and RD) were created, and whole-brain probabilistic tractography was performed between all 164 cortical regions detailed above, using the MRTrix3 package (www.mrtrix.org). Initially, 100 million streamlines were generated; these were filtered to 20 million streamlines using the spherical-deconvolution informed filtering of tractograms (SIFT) algorithm (Smith et al., 2013) to improve the fit between the FOD and number of streamlines in every voxel.

To reconstruct the IFOF, ILF, UF and CST from the whole brain probabilistic tractography, we identified anterior and posterior cortical termination points for each tract based on previously published reports (Wakana et al., 2007; Pannek et al., 2009; Seo and Jang, 2013; Hau et al., 2016; Latini et al., 2017). Cortical regions from the atlas that corresponded to termination points for each tract were combined to create anterior and posterior ROIs (Table 2). For each tract in each hemisphere, a waypoint of identical size was placed in the same location on each participant's FA image (Table 2 & Fig. 2). To confirm that tracts were reconstructed appropriately, they were qualitatively compared to tracts published in an atlas (Catani and Thiebaut de Schotten, 2012). The total number of streamlines contributing to each reconstructed tract, as provided by MRTrix3, were recorded.

To obtain mean FA and RD along the reconstructed tracts, each tract was made into a mask that was then binarized (thresholded at 0.01), and multiplied by the individual's FA and RD images. We limited our analyses to FA and RD, as these indices differed most robustly between patients and healthy controls in our previous study (Moxon-Emre et al.,

Table 2
Regions of interest (ROIs) from the automated cortical and subcortical parcellation that were combined to create custom anterior, posterior, and waypoint ROIs to reconstruct the ILF, IFOF, UF and CST from the whole-brain probabilistic tractography.

Tract	Anterior ROI	Posterior ROI	Waypoints/exclusions
	Combination of ROIs from automated cortical and subcortical parcellation		
IFOF	Inferior frontal gyrus (pars opercularis, pars triangularis and pars orbitalis) Middle frontal gyrus and sulcus Orbital gyri and sulci Fronto-marginal gyrus and sulcus Transverse frontopolar gyri and sulci	Inferior occipital gyrus and sulcus Middle occipital gyrus Lingual gyrus Superior parietal lobule Snterior transverse temporal gyrus Temporal plane of the superior temporal gyrus Occipital pole Cuneus Lateral occipito-temporal gyrus (fusiform gyrus) Angular gyrus	A rectangular ROI was placed over the most ventral portion of the external capsule (spanning 5 axial, 10 coronal, and 10 sagittal slices; volume = 500 mm^3)
ILF	Anterior transverse temporal gyrus Superior temporal gyrus (lateral aspect, planum polare and temporal plane) Middle temporal gyrus Inferior temporal gyrus Parahippocampal gyrus Temporal pole	Superior occipital gyrus Middle occipital gyrus Lateral occipito-temporal gyrus (fusiform gyrus) Cuneus Lingual gyrus Occipital pole	The first slice of the rectangular ROI was placed in the most posterior coronal slice where the temporal lobe was not attached to the frontal lobe (spanning 5 coronal, 15 sagittal, and 15 axial slices; volume = 1125 mm^3)
UF	Inferior frontal gyrus (pars opercularis, pars triangularis and pars orbitalis) Orbital gyri and sulci Straight gyrus Suborbital sulcus Fronto-marginal gyrus Transverse frontopolar gyri and sulci	Temporal pole Superior temporal gyrus (lateral aspect, planum polare) Inferior temporal gyrus Middle temporal gyrus Parahippocampal gyrus Lateral occipito-temporal gyrus (fusiform gyrus) Amgydala	A rectangular ROI was placed where the 'elbow' of the UF is located (spanning 5 coronal, 10 axial and 10 sagittal slices; volume = 500 mm^3). An exclusion mask ROI that covered the entire cerebral hemisphere was placed on the first coronal slice posterior to the amygdala
CST (control tract)	Precentral gyrus		A rectangular ROI was placed over the cerebral peduncle (spanning 5 coronal, 15 axial and 15 sagittal slices; volume = 1125 mm^3). Two exclusion masks were created: 1) A single ROI that covered the entire cerebral hemisphere was placed in the midline, in the sagittal place 2) A single ROI that covered the entire brainstem was placed in the axial view

2016). Examples of the reconstructed tracts, from a single healthy control participant, are shown in Fig. 2.

2.5.4. Tract based spatial statistics (TBSS)

Voxelwise analyses were conducted with TBSS (Smith et al., 2006). All participants' FA data were aligned into a common space (MNI152; Montreal Neurological Institute, McGill, Canada) using the nonlinear registration tool FNIRT (Andersson et al., 2007a, 2007b). Then, a cross-subject mean FA image was created and used to generate a 'skeleton' FA map representing the center common to all tracts, thresholded at FA $>$ 0.20. Finally, participant-specific FA and RD maps were aligned with the skeleton, and values along the width of each tract were considered in the cross-subject voxelwise statistics.

2.6. General cognitive function

2.6.1. Theory of mind (ToM)

Two measures were used to assess ToM. (i) The ability to perspective-take in emotional contexts was measured with a shortened (10-question; score range: 0–40) version of the 25-question Emotional and Emotive Faces task (EEFT) (Dennis et al., 1998, 2013). In this task, participants were read short narratives about a child, and were then asked to indicate: 1) which emotion the child in the story actually feels (emotion identification), and 2) which emotion the child in the story would deliberately choose to express socially (emotive communication). Participants selected the emotion from a board containing 11 drawings of 6 emotions (happy, sad, angry, scared, yucky and neutral), of varying intensity (e.g., very happy, a little bit happy), and this yielded a Feel Inside score (how well they identified the real emotion), Look on Face (how well they identified the concealed emotion) and Concealment (for correctly identifying the reason for concealing the emotion). This task evaluates a child's understanding of, and ability to distinguish between, real emotions, and emotions that are expressed for social purposes. (ii) A parent questionnaire designed to examine a child's ToM capabilities by proxy, the ToM Inventory (ToMI), consisting of 42 items (score range: 0–840; Cronbach's $\alpha = 0.98$) (Hutchins et al., 2011), was also used. This measure yields three subscales: early, basic and advanced ToM. These subscales capture a child's ability to read affect and share attention (early), make use of mental representations and acknowledge them as such (basic), and use complex recursion and to understand that the mind is an active interpreter (advanced) (Hutchins et al., 2011).

2.6.2. Intelligence

The vocabulary and matrix reasoning subtests from the Weschler Abbreviated Scale of Intelligence (WASI-II) (Wechsler, 2011) were used to obtain an estimate of intellectual functioning (IQ).

2.7. Social functioning

The Child Behaviour Checklist (CBCL) (Achenbach, 1991), a parent-report measure designed to evaluate a child's social, academic,

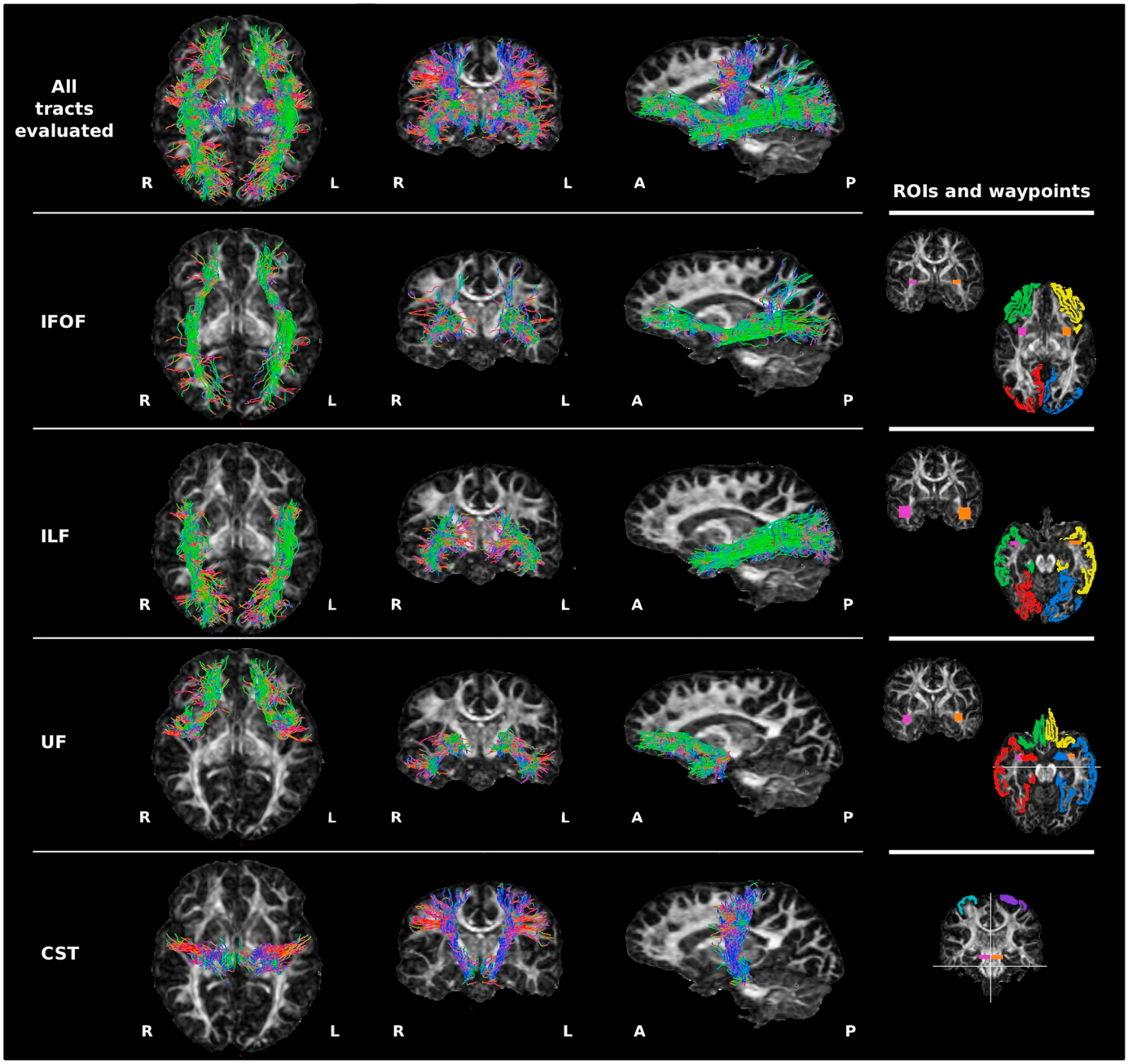


Fig. 2. Examples of the reconstructed WM tracts from a single healthy control participant. Streamlines are overlaid on the FA map and shown in axial, coronal and sagittal planes for: all WM tracts shown together, the inferior frontal occipital fasciculus (IFOF), inferior longitudinal fasciculus (ILF), uncinate fasciculus (UF), and corticospinal tract (CST). Colors of the streamlines represent the fibre orientations: red = medial-lateral; green = anterior-posterior; blue = dorsal-ventral. *Right panel:* the anterior, posterior and waypoint regions of interest (ROIs) used to reconstruct these tracts from the whole-brain probabilistic tractography. White lines indicate exclusion mask placement. For the IFOF, ILF, and UF: green = right anterior ROI; yellow = left anterior ROI; red = right posterior ROI; blue = left posterior ROI. For the CST: blue = right ROI; purple = left ROI. For all tracts: purple = right waypoint; orange = left waypoint. L = left; R = right; A = anterior; P = posterior. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

behavioural and emotional functioning, was used. The social problems subscale from the CBCL consisted of 11 items (score range: 0–22; Cronbach's $\alpha = 0.82$) (Achenbach and Rescorla, 2001) and was utilized to capture social functioning in the current study. The parent version of the Conners-3 (Conners, 2008) to assess ADHD symptoms; it's peer relations subscale consisted of 6 items (score range: 0–18; Cronbach's $\alpha = 0.85$) (Gallant et al., 2007) and was used to capture an additional aspect of social functioning.

2.8. Analytic plan

First, we examined how our measures of visual attention, brain structure and general cognitive function differed between our typically developing and clinical samples. We then tested a model to evaluate the relations among these measures, and examined if they predicted FER. Our primary analyses were conducted to test the hypothesis that

treatment for PF tumours would be associated with damage to the ILF, IFOF and UF, and that this would negatively influence FER abilities. In this model we also evaluated if disrupted visual attention and general cognitive function negatively influenced FER abilities. Our secondary analyses were exploratory; we tested variations on the same model to evaluate if WM damage more broadly (i.e., on a voxelwise basis), or damage to the cerebellum (i.e., cerebellar grey and WM volumes), instead of the abovementioned WM tracts, or contribute to patients' FER deficit. Age was included as a covariate in analyses for all measures that did not yield age-normalized scores. All analyses were corrected for multiple comparisons with Bonferroni correction.

2.8.1. FER

Number of errors on the DANVA2 task were compared between healthy control, surgery and radiation groups using an analysis of variance (ANOVA), controlling for age. Performance on the DANVA2 task is herein referred to as FER. The following are factors that may contribute to FER:

2.8.1.1. Visual attention. We compared the number of fixations, and total time spent looking at the photograph (i.e., total dwell time) during the FER task, between healthy control, surgery and radiation groups using a multivariate ANOVA (MANOVA), as these eye-tracking metrics are non-orthogonal. The number of fixations and dwell time on trials judged correctly vs. incorrectly, within and between groups, were evaluated with repeated measures ANOVAs; this analysis was conducted solely to investigate if the photographs judged incorrectly were viewed differently from those judged correctly, both within and across groups. Age was included as a covariate in all analyses.

2.8.1.2. Brain structure

2.8.1.2.1. Planned analysis. ***Probabilistic Tractography***: For each reconstructed WM tract (ILF, IFOF, UF and CST), two MANOVAs were conducted to compare: i) FA and RD in each hemisphere, and ii) streamline counts in each hemisphere, between healthy control, surgery and radiation groups, controlling for age. In each group, a series of partial correlations were conducted between FA/RD and FER, for each WM tract in each hemisphere, controlling for age. Given that 16 separate correlations were conducted, results were false discovery rate (FDR) corrected at q = 0.1.

2.8.1.2.2. Exploratory analyses. ***Voxelwise analyses*****:** TBSS controls for family-wise errors using a permutation methodology. The null distribution of the cluster-size statistic was built up over 5000 random permutations. Cluster size was thresholded at $p < .05$, which is fully corrected for multiple comparisons. First, we evaluated voxels that differed between healthy controls and all patients considered together, as well as between the healthy control, surgery and radiation groups considered separately. Next, we assessed if any voxels throughout the brain correlated with FER, in healthy control and all patients considered together, and between healthy control and patient groups considered separately. Age was included as a covariate in all analyses. A mask was made for each significant cluster of > 100 voxels, and the anatomic extent of each was labelled with reference to the JHU White-Matter Tractography Atlases (Hua et al., 2008).

Cerebellar volume: grey and WM cerebellar volumes in each hemisphere, normalized to intracranial volume (ICV) (i.e., divided by the total ICV and multiplied by 100), were compared between healthy control, surgery and radiation groups using a MANOVA, controlling for age.

2.8.1.3. General cognitive function. ***Intelligence*****:** Age-standardized scores from the WASI-II (2-subtest IQ) were compared between the healthy control, surgery and radiation groups using an ANOVA.

Theory of Mind**:** Performance on the ToM task (EEFT and ToMI) measures were compared between the healthy control, surgery and radiation groups using a MANOVA, controlling for age. A MANOVA was used because several scores across and within these two ToM measures were significantly correlated, at $r > 0.27$, $p < .05$, as follows: Feel Inside was correlated with Concealment, EEFT Total Score, Basic ToM and Advanced ToM; Look on Face was correlated with EEFT Total Score; Concealment was correlated with EEFT Total Score; EEFT was correlated with Advanced ToM; Early ToM was correlated with Advanced ToM; Basic ToM was correlated with Advanced ToM.

2.8.2. Social functioning

Age-standardized scores from the CBCL (the social problems subscale) and Conners-3 (peer relations subscale) were compared between the healthy control, surgery and radiation groups, using a MANOVA, because scores from these two measures were significantly correlated, $r = 0.62$, $p < .001$.

In order to extend our analyses from descriptive to that of causality, we conducted PLS path modeling, as follows:

2.8.3. Model testing

PLS Path Modeling was performed in R (version 3.3.2), using PLSpm, with 5000 bootstraps (Sanchez, 2013). In all models, we began by testing the accuracy of the model as follows: 1) we examined the relationship between the latent constructs and their associated measures (i.e., loading). 2) We assessed how well each measure corresponded to their latent constructs using Dillon–Goldstein's rho. 3) We evaluated the discriminant validity of the model by confirming that cross-loadings for each measure were larger for measures contained in its own latent construct than for cross-loadings with measures belonging to other latent constructs. To assess the quality of the structural model, we evaluated: 1) the significance of the regression paths (*t*-test), 2) R^2 coefficients of the endogenous variables, with values < 0.2 considered low, and values between 0.2 and 0.5 considered moderate, 3) the average variance extracted (AVE) and 4) goodness-of-fit (GoF), the geometric mean of the average communality and average R^2. GoF values > 0.36 were considered a good fit (Tenenhaus et al., 2005), and bootstrap confidence intervals for path weights and R^2 did not contain zero. PLS path modeling cannot accommodate missing data; thus, participants without MRI data, and/or with missing parent-questionnaire data, were excluded from path analyses (final $n = 42$).

In light of our smaller sample size for the PLS analysis, and the heterogeneity of treatment within our surgery and radiation groups, we elected to characterize our participants along a continuum of treatment type, intensity, complications and time (herein referred to as “medical variables”). We used information from 3 of the 4 domains from the Neurological Predictor Scale (NPS) (Micklewright et al., 2008); for each participant, scores on the surgery, chemotherapy and radiation domains were calculated and summed (NPS score range = 0–7). We created a composite score to capture post-surgical details by attributing each of the following a value of 1: presence of hydrocephalus requiring CSF diversion, cerebellar mutism, and any post-surgical complication (i.e., experiencing any one, or any combination of the following, would yield a score of 1: motor deficits, cranial nerve deficit, visual impairment and hearing impairment) (post-surgical details score range = 0–3). Thus, healthy control participants always scored 0, whereas a patient with a score of 10 on the combined NPS and post-surgical detail score would have received maximal therapy, and experienced considerable post-surgical complications; namely, they would have undergone multiple surgeries, received chemotherapy, craniospinal radiation, developed hydrocephalus that required CSF diversion, cerebellar mutism, and had at least one other post-surgical complication. As it is well documented that some deficits experienced by PF tumour patients, such as cognitive problems, become more apparent over time, time since diagnosis was also included as a medical variable (Mulhern et al., 2004).

Model testing, primary analyses: we hypothesized that the medical variables associated with treatment for PF tumours, would have direct and indirect effects on FER via WM tracts known to be involved in FER (IFOF, ILF, UF – *model 1*; Fig. 1B), but not our control tract (CST - *model*

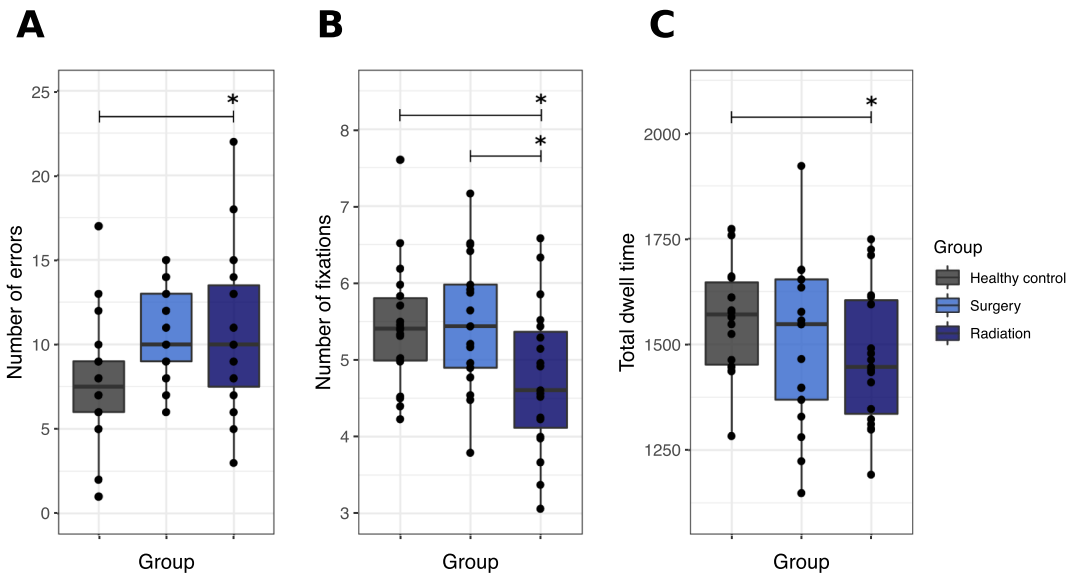


Fig. 3. Behavioural and eye-tracking results from the DANVA2. Boxplots showing all data points with the median (black line) in healthy control, surgery and radiation groups for: A, The total number of FER errors ([Healthy control: Mean = 7.83, SD = 3.78]; [Surgery: Mean = 10.71, SD = 2.82]; [Radiation: Mean = 10.63, SD = 4.78]). B, The total number of fixations made on the photograph ([Healthy control: Mean = 5.44, SD = 0.83]; [Surgery: Mean = 5.49, SD = 0.88]; [Radiation: Mean = 4.75, SD = 0.95]). C, The total time spent looking at the photographs (i.e., total dwell time) ([Healthy control: Mean = 1564.59, SD = 132.82]; [Surgery: Mean = 1501.02, SD = 199.15]; [Radiation: Mean = 1471.23, SD = 160.24]). * $p < .05$.

2; Fig. 1C), and potentially also through its effects on visual attention (i.e., eye-movements), and general cognitive function (i.e., ToM and IQ). The full model structure, and all paths tested, are provided in Fig. 1A (top panel).

Model testing, secondary analyses: The PLS path models were unchanged from the primary analyses except that the WM tracts were replaced with: FA/RD from voxels that differed significantly between patients and healthy controls in the 2-group TBSS analysis (*model 3*; Fig. 1D), and cerebellar grey and WM volumes (*model 4*; Fig. 1E). The full model structure, and all paths tested, are provided in Fig. 1A (bottom panel).

3. Results

3.1. FER

Patients make more errors than healthy controls. Errors on the DANVA2 task differed between healthy controls (mean = 7.58, SD = 3.78) and patients (mean = 10.67, SD = 3.92), ($F(1,51) = 8.11$, $p < .01$, $\eta_p^2 = 0.14$). The three groups differed in FER errors on the DANVA2 task ($F(2,50) = 4.03$, $p = .02$, $\eta_p^2 = 0.14$; Fig. 3A). Both the radiation and surgery groups made more errors than the healthy control group; the radiation group differed significantly ($p = .04$, [CI_{95}: −6.69, −0.13]; Fig. 3A) from controls, whereas the surgery group approached significance ($p = .08$, [CI_{95}: −0.22, 6.26]; Fig. 3A). The following are factors that may contribute to FER:

3.1.1. Visual attention

The radiation group processes the photographs less extensively, but all groups appear to be attending to the photographs. The groups differed in visual attention (number of fixations and total dwell time) ($F(2,50) = 3.67$, $p < .01$, $\eta_p^2 = 0.14$). The radiation group made fewer fixations on the photographs than the healthy control group ($p = .03$, [CI_{95}: 0.07, 1.57]; Fig. 3B) and surgery group ($p = .02$, [CI_{95}: 0.09, 1.58]; Fig. 3B). The radiation group also had a lower total dwell time on the photographs than the healthy control group ($p = .02$, [CI_{95}: 19.33, 272.99]; Fig. 3C), but not the surgery group ($p = .41$, [CI_{95}: −48.95, 201.42]; Fig. 3C). The number of fixations and total dwell time did not differ when viewing photographs that were judged incorrectly vs. correctly, in any group (all $F(2,50) < 0.41$, all $p > .05$, all $\eta_p^2 < 0.02$); Fig. 4A-B). Overall, all participants made more fixations on trials that were judged incorrectly vs. correctly ($p = .05$, $\eta_p^2 = 0.08$), even though the total dwell time did not differ ($p = .13$, $\eta_p^2 = 0.05$). Heat maps provide a summary of the fixations made across participants in each group, for each trial; visual inspection of these heat maps revealed that participants in all groups spent most of their time attending to the face when judging the emotion depicted in the photograph (one trial is shown in Fig. 4C).

3.1.2. Brain structure

3.1.2.1. Planned analyses

3.1.2.1.1. Probabilistic tractography. **The IFOF, ILF, UF and CST do not differ between healthy control and patient groups.** FA and RD did not differ between healthy control and patients groups, in either hemisphere, in any WM tract evaluated (all $F(4,39) < 1.93$, all $p > .05$, all $\eta_p^2 < 0.09$; Table 3), and neither did streamline count (all $F(2,41) < 3.14$, all $p > .05$, all $\eta_p^2 < 0.12$; Table 3).

The left UF correlates with FER in healthy controls only. After FDR correction, FA and RD in the left UF remained significantly correlated with FER in healthy controls only (all $r(13) > 0.63$, all $p < .05$, all $q < 0.1$; Fig. 5), indicating that higher FA/lower RD is associated with fewer FER errors. Notably, FER was not correlated with FA or RD of the CST. Regardless of whether patients were considered as a single group, or separated by their treatment, there were no significant correlations between FA or RD of any tract and FER, before or after FDR correction (all $r(25) < 0.3$, all $p > .05$, Fig. 5; all patients are considered together).

3.1.2.2. Exploratory analyses

3.1.2.2.1. Voxelwise analyses. **Both patient groups differ from healthy controls, whereas the patient groups do not differ from each other.** Despite not differing in the tracts examined, FA and RD differed between healthy control and patient groups (both when considered together, and separated by their treatment), in many clusters of voxels throughout the entire brain (all $F > 8$; all $p < .05$). Post-hoc pairwise comparisons revealed many voxels where, compared to the healthy control group, the surgery group had lower FA and higher RD (all $T > 1.85$; all $p < .05$; Fig. 6A&C,

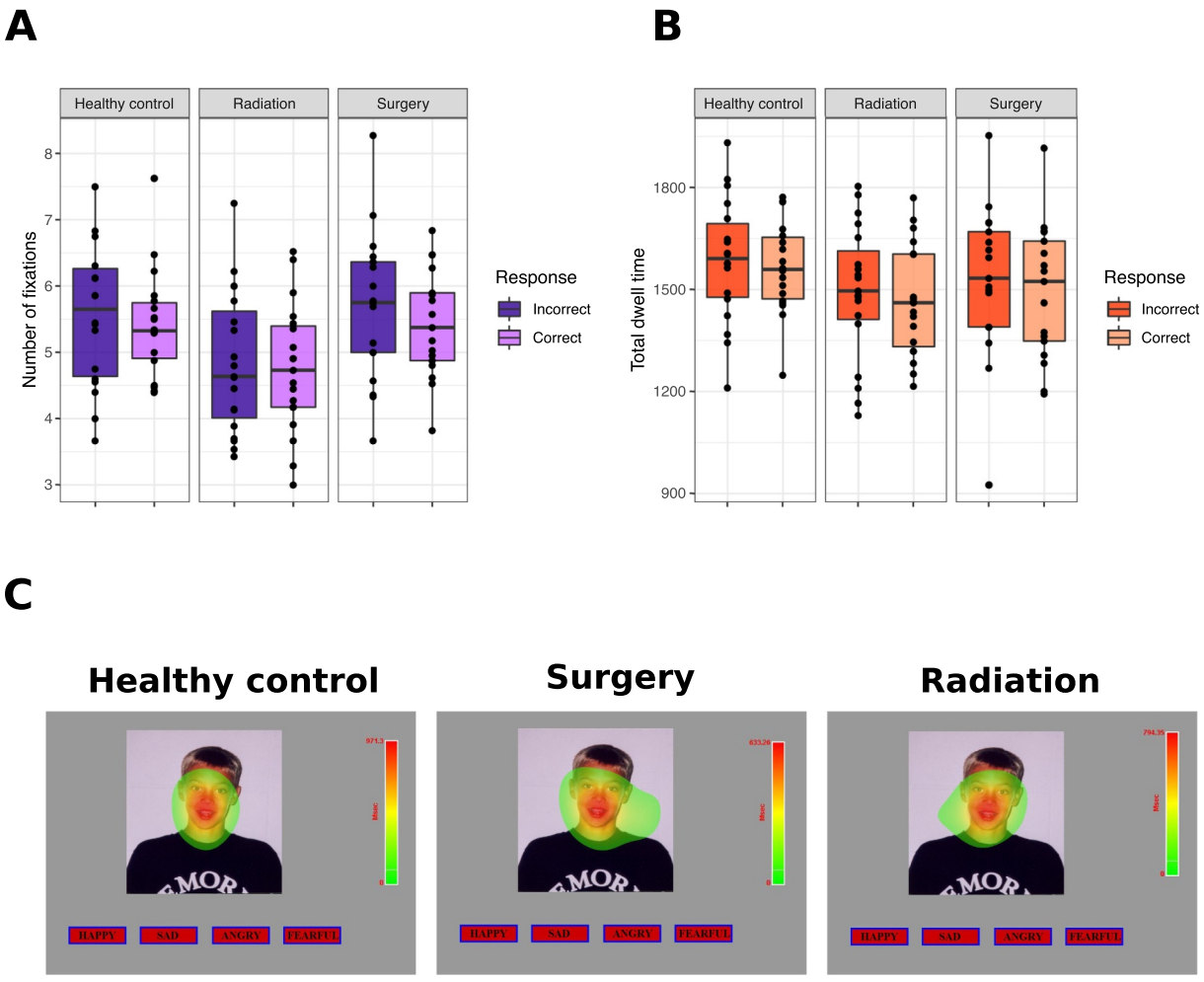


Fig. 4. Eye-tracking results from the DANVA2. A-B. Boxplots showing all data points with the median (black line) in healthy control, surgery and radiation groups for: A, The total number of fixations made on the photograph on correct and incorrect trials. B, The total time spent looking at the photographs (i.e., total dwell time) on correct and incorrect trials. C. Heat maps summarizing the fixations made across all individuals in each group, on a single DANVA2 trial. Warmer colors reflect longer fixations made at that location. The upper limit of the heat map legend reflects the longest fixation made, and this was unique to each group. Visual inspection reveals that individuals in all three groups spent most of their time attending to the face when judging the emotion depicted in the photograph. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4), and where the radiation group had lower FA and higher RD (all $T > 1.74$; all $p < .05$; Fig. 6B&D, Table 4). There were no voxels in either patient group where FA was higher, and RD was lower, than healthy controls. Furthermore, FA and RD did not differ in any voxels between the surgery and radiation groups (all p > .05; Fig. 6E Table 4).

RD in many voxels correlates with FER in healthy controls only. RD, but not FA, was significantly correlated with FER in many voxels throughout the brain, in healthy controls only (all $T > 1.72$; all $p < .05$; Fig. 7, Table 5). Notably, voxels located along the ILF, IFOF and UF (i.e., tracts that we hypothesized would be involved in FER) correlated with performance; however, voxels in many other regions throughout the brain also correlated with performance, the CST (i.e., our control tract) included.

3.1.2.2.2. Cerebellar volume. **Given that patients had their PF tumours surgically removed, patients have smaller cerebellar grey and WM volumes than healthy controls.** The groups differed in cerebellar grey and WM volumes ($F(4,40) = 7.57$, $p < .001$, $\eta_p^2 = 0.34$; Table 3). In the right hemisphere, the radiation and surgery groups had smaller grey and WM volumes than the healthy control group (all p < .05; Table 3). In the left hemisphere, the radiation group had smaller WM volumes than both the surgery and healthy control groups (all $p < .05$; Table 3), and smaller grey matter volume than healthy controls only (all p < .05; Table 3).

3.1.3. General cognitive function

Both patient groups have lower IQ than healthy controls. The effect of group on IQ was significant ($F(2,50) = 6.97$, $p < .01$, $\eta_p^2 = 0.22$; Table 6); post-hoc analyses revealed that the healthy control group had higher full scale IQ than both patient groups (all p < .05; Table 6). It is notable, however, that both patient groups performed very close to the normative mean of 100 (SD = 15). Given that 56% of our controls were siblings of brain tumour patients, this may reflect a true difference in functioning, rather than one driven by the inclusion of controls that may not necessarily represent the general population (i.e., children of clinicians and hospital research staff).

The radiation group has more difficultly with perspective taking in emotional contexts than healthy controls, but parent reported ToM does not highlight any group differences. The MANOVA that included all ToM scores did not yield a significant effect of group ($F(14,82) = 1.50$; $p = .13$, $\eta_p^2 = 0.20$). However, the univariate analyses for each measure revealed that the healthy control and patient groups differed in the Concealment ($F(2,46) = 4.6$ $p = .02$, $\eta_p^2 = 0.17$; Table 6) and Total Scores ($F(2,46) = 4.84$ $p = .01$,

Table 3
FA, RD and streamline count for the left and right ILF, IFOF, UF, CST, in addition to grey and WM volumes of the cerebellum, in healthy control, surgery and radiation groups.

		Healthy control	Surgery	Radiation	F value	p value	η_p^2
IFOF							
FA	Left	0.474 (0.019)	0.469 (0.021)	0.474 (0.020)	0.34	0.71	0.02
	Right	0.477 (0.020)	0.475 (0.021)	0.477 (0.025)	0.09	0.91	< 0.01
RD	Left	0.00058 (0.00002)	0.00060 (0.00004)	0.00059 (0.00004)	1.56	0.22	0.07
	Right	0.00058 (0.00003)	0.00060 (0.00003)	0.00058 (0.00005)	1.32	0.28	0.06
Streamline count	Left	22.12 (15.374)	21.46 (16.071)	21.93 (18.258)	0.02	0.98	< 0.01
	Right	26.71 (24.443)	27.62 (14.327)	32.20 (21.492)	0.09	0.92	< 0.01
ILF							
FA	Left	0.427 (0.027)	0.424 (0.021)	0.430 (0.029)	0.19	0.83	< 0.01
	Right	0.432 (0.024)	0.426 (0.027)	0.430 (0.029)	0.33	0.72	0.02
RD	Left	0.00062 (0.00003)	0.00064 (0.00003)	0.00063 (0.00005)	1.49	0.24	0.07
	Right	0.00062 (0.00003)	0.00064 (0.00004)	0.00063 (0.00005)	1.92	0.16	0.09
Streamline count	Left	195.88 (112.055)	129.31 (60.289)	122.20 (60.521)	3.13	0.05	0.12
	Right	189.12 (118.137)	145.54 (54.428)	184.07 (109.852)	0.76	0.47	0.04
UF							
FA	Left	0.409 (0.013)	0.406 (0.020)	0.405 (0.026)	0.18	0.83	< 0.01
	Right	0.394 (0.012)	0.397 (0.029)	0.389 (0.025)	0.65	0.53	0.03
RD	Left	0.00062 (0.00002)	0.00063 (0.00002)	0.00062 (0.00003)	0.83	0.44	0.04
	Right	0.00063 (0.00001)	0.00063 (0.00003)	0.00063 (0.00003)	0.24	0.79	0.01
Streamline count	Left	138.59 (89.348)	130.08 (115.008)	127.67 (101.384)	0.29	0.75	0.01
	Right	346.53 (217.386)	365.77 (156.084)	305.87 (164.237)	0.56	0.58	0.03
CST							
FA	Left	0.512 (0.025)	0.517 (0.030)	0.511 (0.031)	0.40	0.67	0.02
	Right	0.523 (0.031)	0.526 (0.027)	0.526 (0.025)	0.03	0.98	< 0.01
RD	Left	0.00053 (0.00003)	0.00053 (0.00003)	0.00054 (0.00003)	1.49	0.24	0.07
	Right	0.00051 (0.00003)	0.00052 (0.00003)	0.00052 (0.00004)	0.44	0.65	0.02
Streamline count	Left	77.24 (57.169)	70.46 (39.677)	113.53 (91.114)	2.10	0.14	0.09
	Right	91.47 (61.327)	74.38 (45.154)	81.87 (41.578)	0.80	0.46	0.04
Cerebellar volumes							
White Matter	Left	0.84 (0.09)[a]	0.74 (0.11)[b]	0.62 (0.17)[a,b]	12.78	< 0.001	0.38
	Right	0.79 (0.08)[c,d]	0.57 (0.17)[c]	0.61 (0.17)[d]	11.55	< 0.001	0.36
Grey Matter	Left	3.38 (0.32)[e]	3.06 (0.41)	2.87 (0.57)[e]	4.31	0.02	0.17
	Right	3.43 (0.30) [f,g]	2.68 (0.61)[f]	2.90 (0.44)[g]	11.02	< 0.001	0.34

Matching letters in different rows indicate a significant difference ($p < .05$) between groups as follows. Mean: [a] $p < .001$; [b] $p = .02$; [c] $p < .001$; [d] $p < .01$; [e] $p = .02$; [f] $p < .001$; [g] $p < .01$.
Abbreviations: IFOF = inferior frontal occipital fasciculus; ILF = inferior longitudinal fasciculus; UF = uncinate fasciculus; CST = corticospinal tract; FA = fractional anisotropy; RD = radial diffusivity.

$\eta_p^2 = 0.17$; Table 6) on the EEFT. Post-hoc analyses revealed group differences between the radiation group and healthy control group, whereas the surgery group did not differ significantly from the other groups; relative to the healthy control group, the radiation group performed more poorly in the Concealment and Total Score (all $p < .05$). Thus, the radiation group appeared to have difficulty with this task overall, and in particular with identifying the reason an emoter may want to conceal an emotion from an observer (i.e., Concealment score). They may also have had difficulty selecting which emotion an emoter will express given a particular emotional scenario (i.e., Feel Inside score). Healthy control and patient groups did not differ in their early, basic or advanced ToM scores (all $F(2,46) < 0.67$, all $p > .05$; Table 6).

3.2. Social functioning

Our healthy control and patient groups do not differ in their parent reported social functioning. The MANOVA that included scores from both social functioning measures did not yield a significant effect of group ($F(4, 92) = 0.99$; $p = .42$). Univariate analyses revealed that The groups did not differ in their parent-reported social problems ($F(2,46) = 1.12$, $p = .34$, $\eta_p^2 = 0.05$; Table 6) or peer relations scores ($F(2,46) = 1.42$, $p = .25$, $\eta_p^2 = 0.06$; Table 6).

3.3. Missing questionnaire data and outliers

For the ToMI, CBCL and Conners-3, data was missing for 5 patients; the parent-questionnaires were not returned. For analyses of FER, visual attention, brain structure, general cognitive function and social functioning, we identified univariate outliers by visually examining boxplots of the data, and multivariate outliers using Mahalanobis distance. Half of all analyses (one of the two univariate analysis, and six of the twelve multivariate analyses) contained 1–2 outliers; deletion of these 1–2 cases in each analysis did not change the significance tests. Given that no differences were detected when the analyses were conducted with and without outliers, we report the analyses with all cases included.

3.4. Model testing – PLS path modeling

3.4.1. Primary analyses

In our primary model, we tested the accuracy of a measurement model with the following latent constructs (indicators that were included initially are listed in parentheses): participant (age at testing, parental education, sex), medical variables (NPS score, complications score, time since diagnosis), MRI (FA and RD of the IFOF, ILF, UF in each hemisphere; RD values were multiplied by −1 to prevent negative

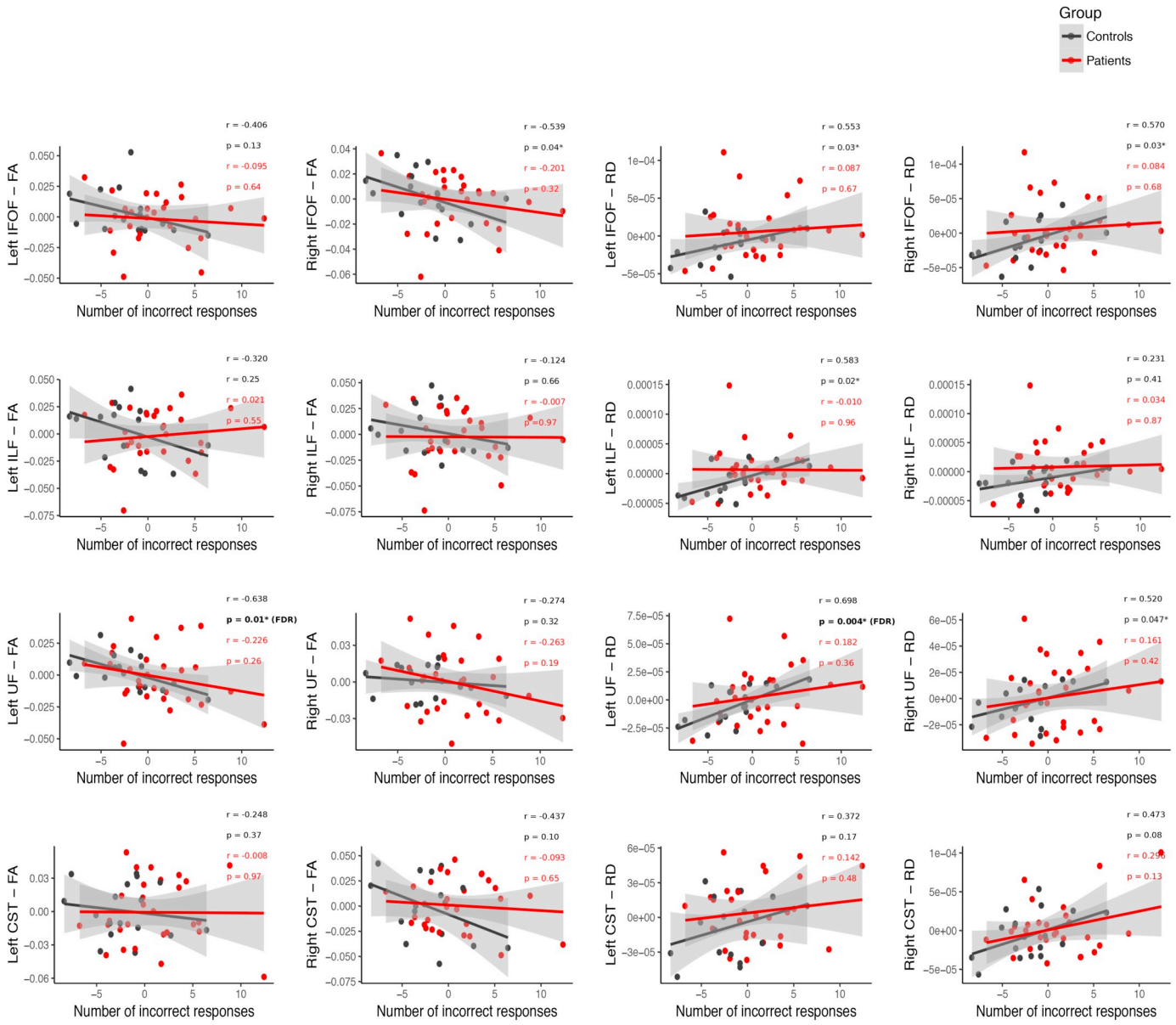
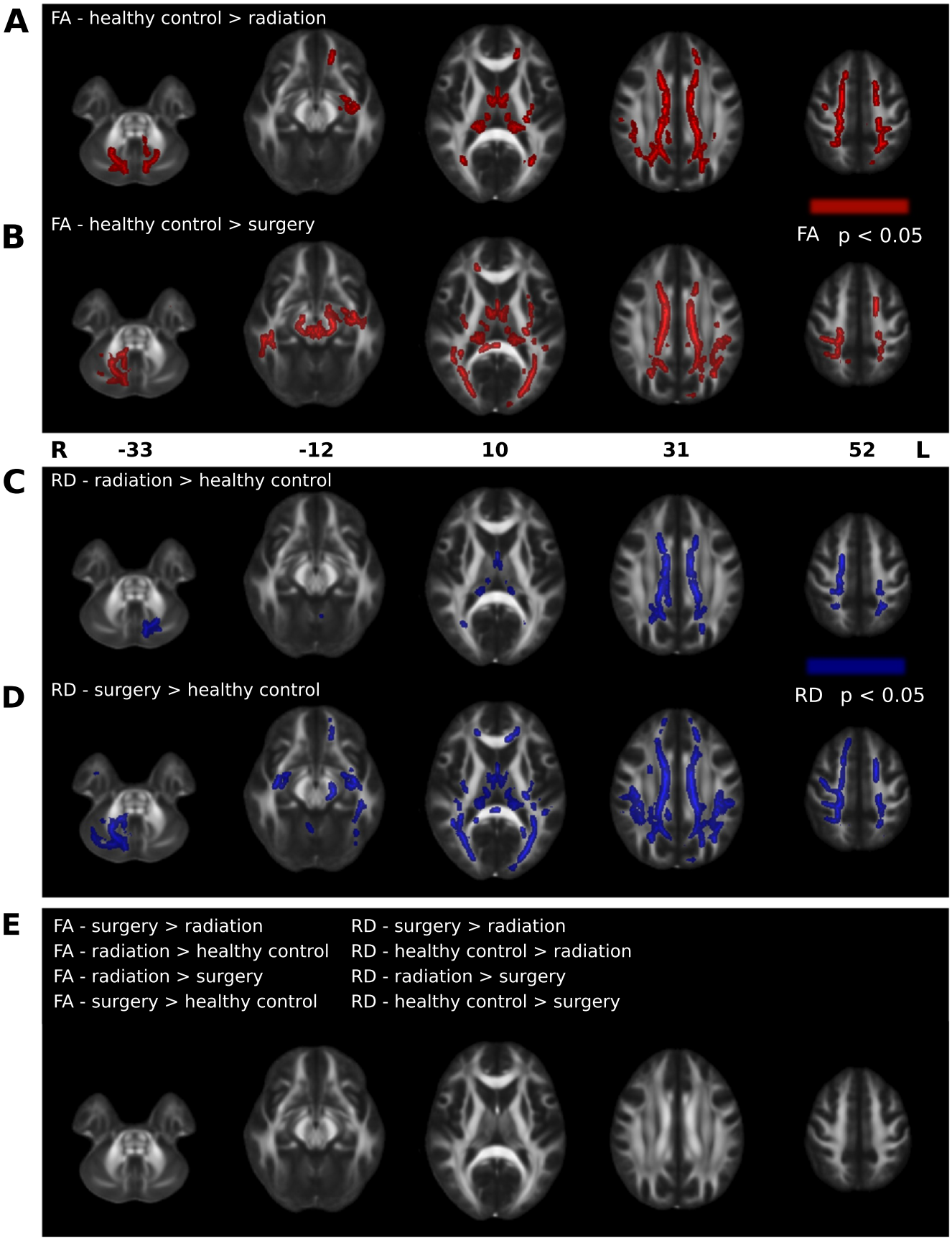


Fig. 5. Partial correlations between FER errors from the DANVA2 task, and FA/RD of the left and right IFOF, ILF, UF and CST, after controlling for age. Partial correlations for healthy controls (black) and patients (red) were conducted separately, but are plotted together to facilitate visual comparison. *Significant correlations ($p < .05$); however, only those indicated with (FDR) survived correction for multiple comparisons ($q < 0.1$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

loadings), IQ (WASI-II score), ToM (early, basic and advanced scores from ToMI; feel inside, concealment and total scores on the EEFT), eye-tracking (fixation count and total dwell time), facial emotion errors (number of incorrect responses on the DANVA2) (Fig. 1A). In our control model, all latent constructs remained unchanged, except that we replaced the MRI indicators with FA and RD of the CST (i.e., our control tract) in each hemisphere. All indicators for both models loaded at > 0.7 on their latent constructs, except for parental education, sex, early ToMI score, basic ToMI score, the EEFT concealment score, and FA of the UF in both hemispheres. These indicators were removed from the model as a result. Thus, the participant latent construct reflects age at testing only, and is herein referred to as age. With all remaining indicators, the latent constructs were all homogenous and unidimensional; Dillon-Goldstein's rho for all latent constructs were > 0.78. In addition, each indicator had a higher cross-loading with the construct it was intended to measure, than with the other latent constructs.

Model 1 (Fig. 1B): medical variables predicted facial emotion errors [$\beta = 0.47$ (CI_{95}: 0.08, 0.88), $t = 2.4$, $p = .02$)] and IQ [$\beta = -0.46$ (CI_{95}: −0.69, −0.21), $t = -3.49$, $p < .01$)]. We also observed that age predicted WM of the ILF, IFOF and UF [$\beta = 0.40$ (CI_{95}: 0.17, 0.67), $t = 2.54$, p = .02)] and ToM [$\beta = 0.54$ (CI_{95}: 0.24, 0.79), $t = 4.08$, $p < .001$)] and WM of the ILF, IFOF and UF predicted IQ [$\beta = 0.29$ (CI_{95}: 0.08, 0.52), $t = 2.2$, $p = .03$)]. No other paths reached statistical significance. The coefficients of determination (R^2) for each latent construct were: WM of the ILF, IFOF and UF = 0.15 (CI_{95}: 0.05, 0.42), IQ = 0.33 (CI_{95}: 0.14, 0.57), eye-tracking = 0.19 (CI_{95}: 0.09, 0.53), ToM = 0.56 (CI_{95}: 0.38, 0.79) and facial emotion errors = 0.33 (CI_{95}: 0.20, 0.64). Model 1 had a GoF of 0.47.

Model 2 (Fig. 1C): medical variables predicted facial emotion errors [$\beta = 0.53$ (CI_{95}: 0.13, 0.89), $t = 2.85$, $p < .01$)] and IQ [$\beta = -0.48$ (CI_{95}: −0.68, −0.25), $t = -3.58$, $p < .001$)]. We also observed that age predicted ToM [$\beta = 0.60$ (CI_{95}: 0.29, 0.84), $t = 4.67$, $p < .0001$)], WM of the CST predicted facial emotion errors [$\beta = -0.35$ (CI_{95}: −0.60, −0.03), $t = -2.51$, $p = .02$)], and IQ predicted ToM



(caption on next page)

Fig. 6. Differences in FA and RD between healthy control and patient groups are shown. A-B, Clusters of voxels with significantly reduced FA (red; p < .05) in the radiation (*A*) and surgery (*B*) groups, compared to healthy controls. C-D, Clusters of voxels with significantly higher RD (blue; p < .05) in the radiation (*C*) and surgery (*D*) groups, compared to healthy controls. E, Comparisons where FA and RD did not differ between groups. Clusters of significant voxels are superimposed on the FMRIB FA template. Images are shown in radiological convention. Numbers represent MNI z-coordinates. Cluster details are provided in Table 4. L = left; R = right. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

[β = −0.35 (CI_{95}: 0.04, 0.61), *t* = 2.46, p = .02)]. No other paths reached statistical significance. The coefficients of determination (R^2) for each latent construct were: WM of the CST = 0.05 (CI_{95}: 0.01, 0.26), IQ = 0.31 (CI_{95}: 0.13, 0.55), eye-tracking = 0.22 (CI_{95}: 0.10, 0.47), ToM = 0.55 (CI_{95}: 0.37, 0.77) and facial emotion errors = 0.42 (CI_{95}: 0.26, 0.66). Model 2 had a GoF of 0.48.

3.4.2. Secondary analyses

In our secondary models, all latent constructs remained unchanged from the original PLS model detailed above, except for the MRI indicators, as follows: model 3 = FA and RD from voxels that differed significantly between the healthy control and patient groups in the TBSS analysis; model 4 = cerebellar grey and WM volumes in each hemisphere.

Model 3 (Fig. 1D): medical variables predicted FA/RD from voxels that differed between patients and healthy controls [β = −0.53 (CI_{95}: −0.78, −0.28), *t* = −3.75, *p* < .001)]. We also observed that age predicted ToM [β = 0.52 (CI_{95}: 0.25, 0.73), *t* = 4.29, p < .001)] and FA/RD from voxels that differed between patients and healthy controls [β = 0.27 (CI_{95}: 0.10, 0.47), *t* = 1.95, *p* = .06), became significant after bootstrapping]. IQ predicted ToM [β = 0.30 (CI_{95}: 0.04, 0.51), *t* = 2.32, p = .03)], and FA/RD from voxels that differed between patients and healthy controls predicted IQ [β = 0.29 (CI_{95}: 0.03, 0.53), *t* = 1.85, *p* = .07), became significant after bootstrapping]. No other paths reached statistical significance. The coefficients of determination (R^2) for each latent construct were: FA/RD from TBSS = 0.28 (CI_{95}: 0.09, 0.60), IQ = 0.25 (CI_{95}: 0.07, 0.51), eye-tracking = 0.20 (CI_{95}: 0.08, 0.51), ToM = 0.58 (CI_{95}: 0.44, 0.80) and facial emotion errors = 0.30 (CI_{95}: 0.22, 0.59). Model 4 had a GoF of 0.50.

Model 4 (Fig. 1E): medical variables predicted facial emotion errors [β = 0.57 (CI_{95}: 0.06, 1.06), *t* = 2.18, *p* = .04)], IQ [β = −0.49 (CI_{95}: −0.79, −0.14), *t* = −2.68, *p* = .01)], and cerebellar volume [β = −0.73 (CI_{95}: −0.88, −0.54), *t* = −5.86, p < .0001)]. We also observed that age predicted ToM [β = 0.53 (CI_{95}: 0.23, 0.78), *t* = 3.91, p < .001)], and that IQ predicted ToM [β = 0.36 (CI_{95}: 0.04, 0.58), *t* = 2.50, p = .02)]. No other paths reached statistical significance. The coefficients of determination (R^2) for each latent construct were: cerebellar volume = 0.47 (CI_{95}: 0.32, 0.69), IQ = 0.25 (CI_{95}: 0.12, 0.48), eye-tracking = 0.22 (CI_{95}: 0.09, 0.54), ToM = 0.54 (CI_{95}: 0.39, 0.76) and facial emotion errors = 0.30 (CI_{95}: 0.22, 0.64) Model 3 had a GoF of 0.49.

In all abovementioned PLS models, the average variance extracted (AVE), a measure of the variance that is captured by the latent construct in relation to the variance that results from measurement error, was above the recommended cutoff of 0.5 (Sanchez, 2013). Across all four models tested, medical variables associated with treatment for PF tumours predicted worse FER, and an older age at testing predicted better ToM capabilities. In three models, medical variables predicted lower IQ, and a lower IQ predicted poorer ToM capabilities. Medical variables predicted lower FA/higher RD values in voxels that differed between patient and healthy control groups, and smaller cerebellar volumes, but neither of these structural metrics predicted FER.

3.4.3. Within-group models – rationale

In light of the correlations we observed between multiple WM measures and FER errors in healthy controls only, we tested a series of PLS path models in both healthy controls and patients to evaluate if increasing age predicts better FER abilities through the ILF/IFOF/UF, and not the CST, in healthy controls only.

We tested the following two models in healthy controls (*n* = 16) and patients (*n* = 26) separately. In our *first model*, we tested the accuracy of a measurement model with the following latent constructs (indicators listed in parentheses): age at testing, MRI (FA and RD of the IFOF, ILF, UF in each hemisphere), facial emotion errors (number of incorrect responses on the DANVA2). In our *second model*, the MRI indicators were changed to FA and RD of the CST in each hemisphere. Indicators for all models loaded at > 0.7 on their latent constructs, except for FA of the UF in both hemispheres, and these were removed from the model as a result. With all remaining indicators, the latent constructs were homogenous and unidimensional (Dillon-Goldstein's rho for all latent constructs were > 0.9), and each indicator had a higher cross-loading with the construct it was intended to measure, than with the other latent constructs.

Healthy controls - Model 1: age predicted facial emotion errors [β = −0.40 (CI_{95}: −0.74, −0.04), *t* = −1.88, *p* = .08 – became significant after bootstrapping)], and WM of the ILF, IFOF and UF also predicted facial emotion errors [β = −0.48 (CI_{95}: −0.85, −0.12), *t* = −2.37, *p* = .03)]. No other paths reached statistical significance. The coefficients of determination (R^2) for each latent construct were: WM of the ILF, IFOF and UF = 0.20 (CI_{95}: 0.03, 0.60), FER = 0.57 (CI_{95}: 0.36, 0.85). Model 1 GoF was 0.536.

Healthy controls - Model 2: age predicted facial emotion errors [β = −0.54 (CI_{95}: −0.83, −0.21), *t* = −2.64, *p* = .02)]. No other paths reached statistical significance. The coefficients of determination (R^2) for each latent construct were: WM of the CST = 0.06 (CI_{95}: 0.004, 0.50), FER = 0.48 (CI_{95}: 0.12, 0.83). Model 2 GoF was 0.501.

Patients – Models 1 & 2: No paths reached statistical significance, and the GoF's were 0.217 and 0.191 respectively, suggesting the data did not fit either model.

Together these models demonstrate that increasing age and higher FA/lower RD of the ILF, IFOF and UF, predict better FER in healthy controls only.

4. Discussion

Our study is the first to combine eye-tracking, neuroimaging and cognitive testing in typically developing children and in children treated for PF tumours, to directly investigate three factors that may contribute to FER: visual attention, brain structure and general cognitive function. We demonstrate that WM, and potentially the ILF, IFOF and UF, are associated with FER in typically developing children only. This association was not present in children treated for PF tumours; it may either be altered, or overshadowed by factors associated with their medical condition and treatment. We also demonstrate that although subtle differences in visual attention and general cognitive functioning emerged, these factors were not associated with FER. In light of these findings, we focus our discussion primarily on the relations between WM and FER, and how it manifests differently in patients treated for PF tumours and typically developing children.

Three lines of evidence in our study converged on the finding that WM is associated with FER in typically developing but not children in treated for PF tumours. Although the associations between WM pathways and FER have been investigated in adults (reviewed in Wang et al., 2018), analogous studies are lacking in children. First, we observed that FER correlated with higher FA and lower RD along the left UF, in typically developing children only. Similarly, in typically developing children only, RD in thousands of voxels throughout the brain, anatomically located within but not limited to the ILF, IFOF, UF, and

Table 4
Clusters of voxels that differ between healthy control, surgery and radiation groups.

Contrast	Cluster (number of voxels)	Cluster Family-Wise Error Corrected p	WM structures[a] encompassed	Mean T value
FA				
Radiation < control	11,270	< 0.01	Forceps minor, CST, forceps major, SLF, cingulum (cingulate gyrus), IFOF, ATR, ILF, UF (L), SLF (temporal part), cingulum (hippocampus)	2.02
	3094	0.01	ATR, IFOF, ILF, CST (L), SLF (L), UF (L), SLF (temporal part), forceps minor, cingulum (cingulate gyrus) (R)	2.63
	784	0.04	Forceps minor, ATR (L), IFOF (L), cingulum (cingulate gyrus) (L), UF (L), SLF (L), SLF (temporal part) (L)	1.74
	748	0.02	CST (R), ATR (R)	3.56
	743	0.02	ATR (L), CST	3.91
Radiation < surgery	0	–	–	–
Surgery < control	12,722	0.01	IFOF, ILF, SLF, forceps major and minor, CST, SLF (temporal part), cingulum (cingulate gyrus), ATR, cingulum (hippocampus), UF	1.9
	5197	0.02	ATR, CST, IFOF, ILF, UF (L), SLF, SLF (temporal part), forceps minor, cingulum (hippocampus) (L), cingulum (cingulate gyrus) (R)	2
	2425	0.03	ATR, CST	2.15
	1575	0.04	IFOF (L), forceps major, ILF (L), SLF (L), SLF (temporal part) (L), ATR, cingulum (hippocampus) (L)	2.05
Control < radiation	0	–	–	–
Surgery < radiation	0	–	–	–
Control < Surgery	0	–	–	–
RD				
Radiation < control	0	–	–	–
Radiation < surgery	0	–	–	–
Surgery < control	0	–	–	–
Control < radiation	7109	0.02	CST, forceps major, IFOF, cingulum (cingulate gyrus), SLF, forceps minor, ATR, ILF, cingulum (hippocampus), SLF (temporal part)	2.14
	1229	0.02	CST (L), ATR (L), CST (R)	3.68
	699	0.03	ATR	3.53
Surgery < radiation	0	–	–	–
Control < surgery	19,209	< 0.01	Forceps minor, SLF, forceps major, IFOF, ILF, CST, SLF (temporal part), cingulum (cingulate gyrus), ATR, UF, cingulum (hippocampus)	1.85
	5153	0.03	ATR, CST, IFOF, ILF, UF, SLF, SLF (temporal part), cingulum (hippocampus) (R), forceps minor, cingulum (cingulate gyrus) (R)	1.9
	3500	0.03	CST (R), ATR (R), IFOF (R)	2.27

All structures are listed in order from greatest to least probability of being a member of the labelled regions within the atlas.
Abbreviations: ATR = anterior thalamic radiation; CST = corticospinal tract; IFOF = inferior fronto-occipital fasciculus; ILF = inferior longitudinal fasciculus; SLF = superior longitudinal fasciculus; UF = uncinate fasciculus.
Significant clusters ($p < .05$) containing < 100 voxels were excluded.
[a] WM structures: defined by the JHU White-Matter Tractography Atlas. Unless otherwise specified, structures listed refer to bilateral counterparts.

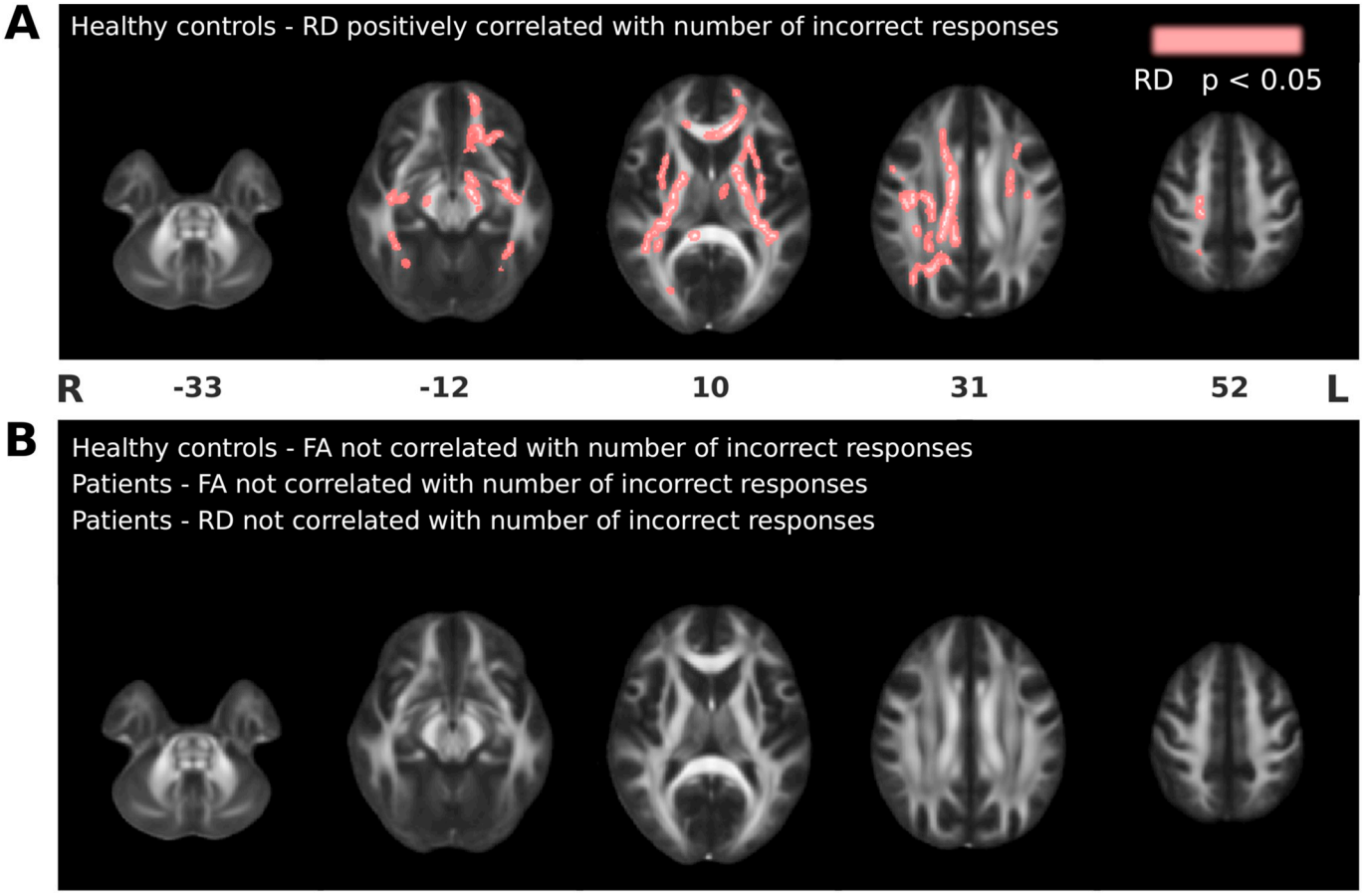


Fig. 7. Correlations between FA/RD and the number of FER errors on the DANVA2. A, Voxels where RD was positively correlated with the number of incorrect responses on the DANVA2 in healthy controls (pink; $p < .05$). B, Correlations in healthy control and patient groups where FA and RD were not correlated with the number of errors on the DANVA2. Clusters of significant voxels are superimposed on the FMRIB FA template. Images are shown in radiological convention. Numbers represent MNI z-coordinates. Cluster details are provided in Table 5. L = left; R = right. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the CST (our control tract), correlated with FER. Lower RD in these voxels predicted fewer errors. Lastly, our PLS path models designed to test the associations between age and FER directly, as well as through their effect on WM (our latent construct measuring FA and RD along the ILF and IFOF bilaterally, and of RD along the UF bilaterally), conducted in typically developing children and children treated for PF tumours separately, revealed that WM predicted FER in typically developing children only. Interestingly, evidence suggests that RD is altered by demyelination and remyelination (Song et al., 2002, 2003, 2005) and RD correlated with FER more consistently than FA. Myelin synthesis is actively occurring during childhood (Lebel and Beaulieu, 2011; Deoni et al., 2012; Dean 3rd et al., 2015); a developmental process that may contribute to the increasing speed and accuracy with which typically developing children recognize emotions as they age (Kolb et al., 1992; De Sonneville et al., 2002).

We confirm that children treated for PF tumours in our sample have difficultly recognizing facial emotions (Bonner et al., 2008) despite attending to the photographs; however, WM indices of the ILF, IFOF and UF did not predict FER despite being structurally intact. Our main hypothesis was therefore not supported. That no structural damage was detected, yet the relationship differed from that of typically developing children, suggests structural preservation of these tracts is not sufficient to support their successful FER. It also suggests there is something unique about the patient brain that is preventing an observable relationship with FER. It is possible that factors related to their medical condition and treatment overshadowed the association between WM and FER. Indeed, medical variables (capturing treatment type and intensity, post-surgical complications, and time since diagnosis) directly predicted FER difficulties, independent of the measures used to assess visual attention, brain structure, and general cognitive function. It appears that aspects of their medical experience, not disentangled in the present study, are contributing more strongly to their FER deficit than WM. This divergence of FER predictors highlights the variable associations between brain structure and behavioural outcomes, and cautions against assuming that normal appearing brain structure will play the same functional role in typically developing and clinical populations.

Knowledge that the same brain structure can have different functions has been highlighted as a fundamental problem about making reverse inferences from functional neuroimaging data in cognitive neuroscience (Poldrack, 2012). A recent study conducted in children with autism demonstrated that networks underlying intelligence differ from typically developing children (Pua et al., 2018). Another study, conducted in children born very preterm, identified an absence of an association between DTI metrics and neurodevelopmental outcome that was present in term born children, but found their WM to be less connected (Young et al., 2018). It is possible that FER relies more heavily on brain networks rather than individual tracts, and that network analyses may have been more sensitive to brain differences that are driving our clinical sample's FER deficit.

Across all participants, we found that age predicted ToM directly. Our primary PLS path model also suggests that the age-dependent improvement of ToM may be partially related to maturation of the ILF, IFOF and UF, through their effect on IQ. Meaning, WM structure (i.e., a

Table 5
Clusters of voxels that correlate with FER (number of incorrect responses on the DANVA2 task).

Contrast	Cluster (number of voxels)	Cluster Family-Wise Error Corrected p	WM structures[a] encompassed	Mean T value
Controls				
FA				
Positive correlation	0	–	–	–
Negative correlation	0	–	–	–
RD				
Positive correlation	5016	0.03	CST, SLF (R), IFOF (R), SLF (temporal part) (R), forceps minor, ILF (R), cingulum (cingulate gyrus) (R), forceps major, ATR, UF (R), cingulum (hippocampus) (R)	1.72
	4315	0.03	Forceps minor, ATR, IFOF (L), CST (L), UF (L), SLF (L), ILF (L), SLF (temporal part), cingulum (cingulate gyrus)	1.88
	692	0.04	Forceps major, IFOF (R), ILF (R), SLF (R), cingulum (hippocampus), ATR (R), cingulum (cingulate gyrus) (R)	1.91
	331	0.04	IFOF (L), ILF (L), SLF (temporal part) (L), SLF (L), UF (L), ATR (L), forceps minor, cingulum (hippocampus) (L)	2.07
	167	0.04	CST (L), SLF, ATR, cingulum (cingulate gyrus) (L), SLF (temporal part) (R)	2.77
	162	0.04	SLF (L), SLF (temporal part) (L)	3.14
Negative correlation	0	–	–	–
Patients				
FA				
Positive correlation	0	–	–	–
Negative correlation	0	–	–	–
RD				
Positive correlation	0	–	–	–
Negative correlation	0	–	–	–

All structures are listed in order from greatest to least probability of being a member of the labelled regions within the atlas.
Abbreviations: ATR = anterior thalamic radiation; CST = corticospinal tract; IFOF = inferior fronto-occipital fasciculus; ILF = inferior longitudinal fasciculus; SLF = superior longitudinal fasciculus; UF = uncinate fasciculus.
Significant clusters ($p < .05$) containing < 100 voxels were excluded.
[a] WM structures: defined by the JHU White-Matter Tractography Atlas. Unless otherwise specified, structures listed refer to bilateral counterparts.

Table 6
Scores of measures used to evaluate general cognitive function, and social functioning scores from parent questionnaires, in healthy control, surgery and radiation groups.

Measure	Healthy control	Surgery	Radiation	F value	p value	η_p^2
WASI-II						
2-subtest IQ	115.72 (10.55)[a,b]	103.35 (17.44)[a]	97.83 (15.46)[b]	6.97	< 0.01	0.22
Theory of Mind Measures - Multivariate Test				1.50	0.13	0.20
EEFT						
Feel Inside	8.17 (1.47)	7.41 (1.62)	7.58 (1.22)	2.66	0.08	0.10
Look on Face	13.17 (3.82)	14.29 (3.98)	12.37 (3.00)	1.23	0.30	0.05
Concealment	8.00 (2.03)[a]	7.29 (2.62)	5.63 (2.93)[a]	4.63	0.02	0.17
Total Score	30.22 (5.34)[b]	29.00 (6.22)	25.58 (4.56)[b]	4.84	0.01	0.17
ToMI						
Early	135.31 (4.39)	134.61 (8.53)	134.29 (7.14)	0.21	0.81	0.009
Basic	361.69 (27.61)	361.69 (19.49)	371.63 (8.99)	0.66	0.52	0.03
Advanced	291.98 (32.11)	290.52 (24.28)	293.09 (23.64)	0.53	0.60	0.02
Social functioning measures - multivariate test				0.99	0.42	0.04
CBCL						
Social Problems	54.17 (4.77)	55.47 (6.09)	58.33 (10.47)	1.12	0.34	0.05
Conners-3						
Peer Relations	52.11 (7.68)	58.88 (15.82)	61.27 (17.60)	1.42	0.25	0.06

Values provided are means (standard deviation). Matching letters in different rows indicate a significant difference ($p < .05$) between groups as follows: [a]$p = .05$; [b]$p < .01$; [c]$p = .01$; [d]$p = .01$.
Abbreviations: EEFT = Emotional and Emotive Faces task; ToMI = Theory of Mind Inventory; WASI-II = Weschler Abbreviated Scale of Intelligence; CBCL = Child Behavior Checklist.

higher FA/lower RD) increases with age, which is associated with a higher IQ, and a higher IQ predicts better ToM performance. Fittingly, a recent study documented that WM maturation in early childhood is associated with ToM development (Wiesmann et al., 2017). Given that children treated for PF tumours did not differ from typically developing children in ToM, that their IQ was within normal limits, and that they attended to the photographs, it is not surprising that our measures of general cognitive function and visual attention were not associated with their FER abilities. However, our study highlights that some of the deficits currently experienced by children treated for PF tumours are subtle (i.e., impairments in FER, despite general cognitive function within normal limits), and that they could be overlooked if not assessed directly. Moreover, it is important to consider that the age range captured in the present study is broad, and that there may be some differences or non-linear alterations related to biological maturation that can influence FER abilities; for instance, there is evidence of a decrement in face processing and voice recognition abilities at the time of sexual maturation (Steinberg, 2005).

It is well documented that children treated for brain tumours, in particular those who received more intensive treatment, experience social difficulties (Schultz et al., 2007; Bonner et al., 2008; Brinkman et al., 2012). Our groups did not differ on parent-reported measures of social functioning; this either suggests the FER deficits experienced by patients are not currently causing social problems in our sample, or the proxy measures we used to assess social functioning lacked the sensitivity to detect their problems. If patients are indeed not currently experiencing social problems, their FER deficit could still have negative implications in the future, in particular as patients enter adulthood, when social dynamics and expectations evolve from that of childhood.

Some limitations to the current study should be noted, and some future directions proposed. First, only a subset of our sample was included in the PLS-path modeling as this technique requires complete data for all included measures. Thus, patients who did not undergo MRI, and patients who did not return questionnaires, were excluded. Second, the abbreviated version of the Wechsler intelligence scale may lack the sensitivity to capture the full range of deficits experienced by patients. For instance, processing speed, a domain particularly affected by treatment with radiation (Mabbott et al., 2008; Scantlebury et al., 2016), is not evaluated as a part of the WASI-II. Thus, the IQ scores for patients in our radiation group in particular, may represent overestimates. Third, the parent-reported measures of social functioning used in the present study may have lacked the sensitivity to capture the full range of social problems experienced by patients, and these measures do not primarily function as measures of social functioning. Notably, the CBCL was designed for use in typically developing children, and its use in chronically ill populations has been criticized (Perrin et al., 1991). However, these measures of social functioning have been frequently and successfully used in the pediatric brain tumour population, and using them in this study was deemed useful for comparison with the existing literature. And lastly, although the present study investigated factors we expected to contribute to FER, it is likely that we did not capture all the relevant factors. Future studies may consider examining WM pathways thought to be involved in emotional and/or visual processing, such as WM pathways that connect the cerebellum to the limbic system (Blatt et al., 2013) and the optic radiations, respectively. It may also be worthwhile to evaluate differences in the speed of neural processing between typically developing children and children treated for a PF tumour during task performance, as children treated with radiation have been shown to exhibit delayed visual latencies in a visual-motor reaction time task (Dockstader et al., 2013). Moreover, given that PF tumour patients appear to be attending to the faces, it may be worthwhile to assess more nuanced indices of visual attention relevant to emotion recognition, such as evaluating which facial features were attended to during the task. Lastly, larger sample sizes and more homogenous treatment groups will likely be required to disentangle how medical variables are contributing to their emotion recognition deficit.

5. Conclusion

In the present study, we found a divergence of FER predictors, with WM predicting FER in typically developing children, and medical variables predicting the deficit experienced by children treated for PF tumours. We found no associations between visual attention, or general cognitive function, and FER. To mitigate the negative impact of FER deficits in children and adolescents, it is important to understand which factors contribute to their deficit, and how; it is clear that further studies are required to disentangle the role of factors we evaluated, and to examine if factors not captured in the present study also contribute. Importantly, our study provides some insight into predictors that may be specific to children treated for PF tumours, and captured a divergence of associations between typically developing and clinical populations; a concept that may be broadly applicable to other neurodevelopmental and clinical populations that experience FER deficits.

Conflict of interests

None.

Funding

This work was supported by the Canadian Institute of Canadian Institute of Health Research (CIHR) (MOP-123537) and the Pediatric Oncology Group of Ontario.

Acknowledgments

The authors would also like to thank: all the children who participated in this study, Dr. Stephen Nowicki for providing the raw images used to re-create an eye-tracking version of the DANVA2 task he developed, and Dr. Margot J. Taylor for critically reviewing the manuscript and providing helpful suggestions.

References

Achenbach, T., 1991. Manual for the CBCL. University of Vermont, Burlington, VT.
Achenbach, T.M., Rescorla, L.A., 2001. Manual for the ASEBA School-Age Forms and Profiles. University of Vermont Research Center for Children, Youth, & Families, Burlington, VT.
Andersson, J.L.R., Jenkinson, M., Smith, S.M., 2007a. Non-linear optimisation. FMRIB technical report TR07JA1 from. www.fmrib.ox.ac.uk/analysis/techrep.
Andersson, J.L.R., Jenkinson, M., Smith, S.M., 2007b. Non-linear registration, aka Spatial normalisation. FMRIB technical report TR07JA2 from. www.fmrib.ox.ac.uk/analysis/techrep.
Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C., 2011. A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 54, 2033–2044.
Basser, P.J., 1995. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. NMR Biomed. 8, 333–344.
Baum, K.M., Nowicki, S., 1998. Perception of emotion: measuring decoding accuracy of adult prosodic cues varying in intensity. J. Nonverbal Behav. 22, 89–108.
Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system – a technical review. NMR Biomed. 15, 435–455.
Beh, S.C., Frohman, T.C., Frohman, E.M., 2017. Cerebellar control of eye movements. J. Neuroophthalmol. 37, 87–98.
Blatt, G.J., Oblak, A.L., Schmahmann, J.D., 2013. Cerebellar Connections with Limbic Circuits: Anatomy and Functional Implications. Springer, Dordrecht.
Bonner, M.J., Hardy, K.K., Willard, V.W., Anthony, K.K., Hood, M., Gururangan, S., 2008. Social functioning and facial expression recognition in survivors of pediatric brain tumors. J. Pediatr. Psychol. 33, 1142–1152.
Brinkman, T.M., Palmer, S.L., Chen, S., Zhang, H., Evankovich, K., Swain, M.A., Bonner, M.J., Janzen, L., Knight, S., Armstrong, C.L., Boyle, R., Gajjar, A., 2012. Parent-reported social outcomes after treatment for pediatric embryonal tumors: a prospective longitudinal study. J. Clin. Oncol. 30, 4134–4140.
Bunford, N., Evans, S.W., Wymbs, F., 2015. ADHD and emotion dysregulation among children and adolescents. Clin. Child. Fam. Psychol. Rev. 18, 185–217.
Calvo, M.G., Lang, P.J., 2004. Gaze patterns when looking at emotional pictures: motivationally biased attention. Motiv. Emot. 28, 221–243.
Catani, M., Thiebaut de Schotten, M., 2012. Atlas of Human Brain Connections. Oxford. Oxford University Press, New York, NY.
Clausi, S., Iacobacci, C., Lupo, M., Olivito, G., Molinari, M., Leggio, M., 2017. The role of the cerebellum in unconscious and conscious processing of emotions: a review. Appl. Sci. 7, 521.
Coad, B.M., Postans, M., Hodgetts, C.J., Muhlert, N., Graham, K.S., Lawrence, A.D., 2017. Structural connections support emotional connections: uncinate fasciculus microstructure is related to the ability to decode facial emotion expressions. Neuropsychologia S0028–3932, 30420–30427.
Collin, L., Bindra, J., Raju, M., Gillberg, C., Minnis, H., 2013. Facial emotion recognition in child psychiatry: a systematic review. Res. Dev. Disabil. 34, 1505–1520.
Conners, K.C., 2008. Conners 3rd Edition. Multi-Health Systems, Toronto, Ontario, Canada.
Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C.Y., Barnett, J.H., 2018. Social cognitive dysfunction as a clinical marker: a systematic review of meta-analyses across 30 clinical conditions. Neurosci. Biobehav. Rev. 84, 92–99.
De Sonneville, L.M., Verschoor, C.A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., Vranken, M., 2002. Facial identity and facial emotions: speed, accuracy, and processing strategies in children and adults. J. Clin. Exp. Neuropsychol. 24, 200–213.
Dean 3rd, D.C., O'Muircheartaigh, J., Dirks, H., Waskiewicz, N., Walker, L., Doernberg, E., Piryatinsky, I., Deoni, S.C., 2015. Characterizing longitudinal white matter development during early childhood. Brain Struct. Funct. 220, 1921–1933.
Dennis, M., Barnes, M.A., Wilkinson, M., Humphreys, R.P., 1998. How children with head injury represent real and deceptive emotion in short narratives. Brain Lang. 61, 450–483.
Dennis, M., Agostino, A., Taylor, H.G., Bigler, E.D., Rubin, K., Vannatta, K., Gerhardt, C.A., Stancin, T., Yeates, K.O., 2013. Emotional expression and socially modulated

emotive communication in children with traumatic brain injury. J. Int. Neuropsychol. Soc. 19, 34–43.
Deoni, S.C., Dean 3rd, D.C., O'Muircheartaigh, J., Dirks, H., Jerskey, B.A., 2012. Investigating white matter development in infancy and early childhood using myelin water faction and relaxation time mapping. Neuroimage 63, 1038–1053.
Destrieux, C., Fischl, B., Dale, A., Halgren, E., 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53, 1–15.
Dockstader, C., Gaetz, W., Bouffet, E., Tabori, U., Wang, F., Bostan, S.R., Laughlin, S., Mabbott, D.J., 2013. Neural correlates of delayed visual-motor performance in children treated for brain tumours. Cortex 49, 2140–2150.
Duc, A.H., Bays, P., Husain, M., 2008. Eye movements as a probe of attention. Prog. Brain Res. 171, 403–411.
Edwards, M., Stewart, E., Palermo, R., Lah, S., 2017. Facial emotion perception in patients with epilepsy: a systematic review with meta-analysis. Neurosci. Biobehav. Rev. 83, 212–225.
Eussen, M.L., Louwerse, A., Herba, C.M., Van Gool, A.R., Verheij, F., Verhulst, F.C., Greaves-Lord, K., 2015. Childhood facial recognition predicts adolescent symptom severity in autism spectrum disorder. Autism Res. 8, 261–271.
Evers, K., Steyaert, J., Noens, I., Wagemans, J., 2015. Reduced recognition of dynamic facial emotional expressions and emotion-specific response bias in children with an autism spectrum disorder. J. Autism Dev. Disord. 45, 1774–1784.
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., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex 14, 11–22.
Gallant, S., Conners, C.K., Rzepa, S.R., Pitkanen, J., Marocco, M., Sitarenious, G., 2007. Psychometric properties of the conners 3. In: Poster Session Presented at the Annual Meeting of the American Psychological Association, (San Francisco, CA).
Hau, J., Sarubbo, S., Perchey, G., Crivello, F., Zago, L., Mellet, E., Jobard, G., Joliot, M., Mazoyer, B.M., Tzourio-Mazoyer, N., Petit, L., 2016. Cortical terminations of the inferior fronto-occipital and *Uncinate fasciculi*: anatomical stem-based virtual dissection. Front. Neuroanat. 10, 58.
Henderson, J.M., Hollingworth, A., 1998. Eye movements during scene viewing: an overview. In: Underwood, G. (Ed.), Eye Guidance in Reading and Scene Perception. Elsevier Science Ltd., Oxford, pp. 269–283.
Herba, C., Phillips, M., 2004. Annotation: development of facial expression recognition from childhood to adolescence: behavioural and neurological perspectives. J. Child Psychol. Psychiatry 45, 1185–1198.
Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., Calabresi, P.A., Pekar, J.J., van Zijl, P.C., Mori, S., 2008. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. Neuroimage 39, 336–347.
Hutchins, T.L., Prelock, P.A., Bonazinga, L., 2011. Psychometric evaluation of the Theory of Mind Inventory (ToMI): a study of typically developing children and children with autism spectrum disorder. J. Autism Dev. Disord. 42, 327–341.
Im-Bolter, N., Agostino, A., Owens-Jaffray, K., 2016. Theory of mind in middle childhood and early adolescence: different from before? J. Exp. Child Psychol. 149, 98–115.
Jones, D.K., Leemans, A., 2011. Diffusion tensor imaging. Methods Mol. Biol. 711, 127–144.
Kolb, B., Wilson, B., Taylor, L., 1992. Developmental changes in the recognition and comprehension of facial expression: implications for frontal lobe function. Brain Cogn. 20, 74–84.
Latini, F., Martensson, J., Larsson, E.M., Fredrikson, M., Ahs, F., Hjortberg, M., Aldskogius, H., Ryttlefors, M., 2017. Segmentation of the inferior longitudinal fasciculus in the human brain: a white matter dissection and diffusion tensor tractography study. Brain Res. 1675, 102–115.
Lawrence, K., Campbell, R., Skuse, D., 2015. Age, gender, and puberty influence the development of facial emotion recognition. Front. Psychol. 6, 761.
Lebel, C., Beaulieu, C., 2011. Longitudinal development of human brain wiring continues from childhood into adulthood. J. Neurosci. 31, 10937–10947.
Lee, S.B., Koo, S.J., Song, Y.Y., Lee, M.K., Jeong, Y.J., Kwon, C., Park, K.R., Park, J.Y., Kang, J.I., Lee, E., An, S.K., 2014. Theory of mind as a mediator of reasoning and facial emotion recognition: findings from 200 healthy people. Psychiatry Investig. 11, 105–111.
Luna, B., Velanova, K., Geier, C.F., 2008. Development of eye-movement control. Brain Cogn. 68, 293–308.
Mabbott, D.J., Penkman, L., Witol, A., Strother, D., Bouffet, E., 2008. Core neurocognitive functions in children treated for posterior fossa tumors. Neuropsychology 22, 159–168.
Mancuso, M., Magnani, N., Cantagallo, A., Rossi, G., Capitani, D., Galletti, V., Cardamone, G., Robertson, I.H., 2015. Emotion recognition impairment in traumatic brain injury compared with schizophrenia spectrum: similar deficits with different origins. J. Nerv. Ment. Dis. 203, 87–95.
Micklewright, J.L., King, T.Z., Morris, R.D., Krawiecki, N., 2008. Quantifying pediatric neuro-oncology risk factors: development of the neurological predictor scale. J. Child Neurol. 23, 455–458.
Morris, R.W., Weickert, C.S., Loughland, C.M., 2009. Emotional face processing in schizophrenia. Curr Opin Psychiatry 22, 140–146.
Moxon-Emre, I., Bouffet, E., Taylor, M.D., Laperriere, N., Sharpe, M.B., Laughlin, S., Bartels, U., Scantlebury, N., Law, N., Malkin, D., Skocic, J., Richard, L., Mabbott, D.J., 2016. Vulnerability of white matter to insult during childhood: evidence from patients treated for medulloblastoma. J Neurosurg Pediatr 18, 29–40.
Mulhern, R.K., Merchant, T.E., Gajjar, A., Reddick, W.E., Kun, L.E., 2004. Late neurocognitive sequelae in survivors of brain tumours in childhood. Lancet Oncol. 5, 399–408.
Padovani, L., Andre, N., Constine, L.S., Muracciole, X., 2012. Neurocognitive function after radiotherapy for paediatric brain tumours. Nat. Rev. Neurol. 8, 578–588.
Pannek, K., Chalk, J.B., Finnigan, S., Rose, S.E., 2009. Dynamic corticospinal white matter connectivity changes during stroke recovery: a diffusion tensor probabilistic tractography study. J. Magn. Reson. Imaging 29, 529–536.
Perrin, E.C., Stein, R.E., Drotar, D., 1991. Cautions in using the child behavior checklist: observations based on research about children with a chronic illness. J. Pediatr. Psychol. 16, 411–421.
Poldrack, R.A., 2012. The future of fMRI in cognitive neuroscience. Neuroimage 62, 1216–1220.
Pua, E.P.K., Malpas, C.B., Bowden, S.C., Seal, M.L., 2018. Different brain networks underlying intelligence in autism spectrum disorders. Hum. Brain Mapp. 39, 3253–3262.
Sanchez, G., 2013. PLS Path Modeling with R. Trowchex Editions, Berkeley.
Scantlebury, N., Bouffet, E., Laughlin, S., Strother, D., McConnell, D., Hukin, J., Fryer, C., Laperriere, N., Montour-Proulx, I., Keene, D., Fleming, A., Jabado, N., Liu, F., Riggs, L., Law, N., Mabbott, D.J., 2016. White matter and information processing speed following treatment with cranial-spinal radiation for pediatric brain tumor. Neuropsychology 30, 425–438.
Schultz, K.A., Ness, K.K., Whitton, J., Recklitis, C., Zebrack, B., Robison, L.L., Zeltzer, L., Mertens, A.C., 2007. Behavioral and social outcomes in adolescent survivors of childhood cancer: a report from the childhood cancer survivor study. J. Clin. Oncol. 25, 3649–3656.
Seo, J.P., Jang, S.H., 2013. Different characteristics of the corticospinal tract according to the cerebral origin: DTI study. AJNR Am. J. Neuroradiol. 34, 1359–1363.
Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487–1505.
Smith, R.E., Tournier, J.D., Calamante, F., Connelly, A., 2013. SIFT: spherical-deconvolution informed filtering of tractograms. Neuroimage 67, 298–312.
Song, S.K., Sun, S.W., Ramsbottom, M.J., Chang, C., Russell, J., Cross, A.H., 2002. Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. Neuroimage 17, 1429–1436.
Song, S.K., Sun, S.W., Ju, W.K., Lin, S.J., Cross, A.H., Neufeld, A.H., 2003. Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. Neuroimage 20, 1714–1722.
Song, S.K., Yoshino, J., Le, T.Q., Lin, S.J., Sun, S.W., Cross, A.H., Armstrong, R.C., 2005. Demyelination increases radial diffusivity in corpus callosum of mouse brain. Neuroimage 26, 132–140.
Steinberg, L., 2005. Cognitive and affective development in adolescence. Trends Cogn. Sci. 9, 69–74.
Taylor, L.J., Maybery, M.T., Grayndler, L., Whitehouse, A.J., 2015. Evidence for shared deficits in identifying emotions from faces and from voices in autism spectrum disorders and specific language impairment. Int J Lang Commun Disord 50, 452–466.
Tenenhaus, M., Vinzi, V.E., Chatelin, Y.M., Lauro, C., 2005. PLS path modeling. Comput. Stat. Data Anal. 48, 159–205.
Tournier, J.D., Calamante, F., Connelly, A., 2013. Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. NMR Biomed. 26, 1775–1786.
Wakana, S., Caprihan, A., Panzenboeck, M.M., Fallon, J.H., Perry, M., Gollub, R.L., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., van Zijl, P., Mori, S., 2007. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 36, 630–644.
Wang, Y., Metoki, A., Alm, K.H., Olson, I.R., 2018. White matter pathways and social cognition. Neurosci. Biobehav. Rev. 90, 350–370.
Wechsler, D., 2011. Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II). NCS Pearson, San Antonio, TX.
Wiesmann, C.G., Schreiber, J., Singer, T., Steinbeis, N., Friederici, A.D., 2017. White matter maturation is associated with the emergence of Theory of Mind in early childhood. Nat. Commun. 8 14692.
Young, J.M., Vandewouw, M.M., Morgan, B.R., Smith, M.L., Sled, J.G., Taylor, M.J., 2018. Altered white matter development in children born very preterm. Brain Struct. Funct. 223, 2129–2141.