PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Towards PErsonalized PRognosis for Children with Traumatic Brain Injury: The PEPR Study Protocol
AUTHORS	Kooper, Cece; Oosterlaan, Jaap; Bruining, Hilgo; Engelen, Marc; Pouwels, Petra; Popma, Arne; van Woensel, Job; Buis, Dennis; Steenweg, Marjan; Hunfeld, Maayke; Königs, Marsh

VERSION 1 – REVIEW

REVIEWER	Shah, Sudhin Weill Cornell Medical College
REVIEW RETURNED	06-Jan-2022

GENERAL COMMENTS	The inability to accurately prognosticate the myriad outcomes following heterogeneous TBI certainly warrants large data-driven studies such as the one proposed. There are, however, several concerns about this current protocol. 1) While the MRI based analyses are to be explored for their additive contributions, is it the expectation that if found to be useful, it would translate to clinical practice? i.e. is access to MRI facilities and the very specific expertise needed for analyses scalable? 2) How will the machine learning approaches translate to real-life implementation? i.e. if validated in this initial sample, how is this expected to be used when a future child is admitted and the clinicians need to predict that child's function? 3) What is the justification for excluding those below 4? 4) What is mild TBI criteria? concussions or complicated mild? 5) Exclusion criteria - why is the study excluding anybody who has'nt recovered to the point that they can participate in outcome assessments? i.e. if the purpose is to predict heterogeneous outcome, why exclude those who might still be recovering but might be in a disorder of consciousness or post-traumatic confusional state etc at the time of the assessments? Why not expand the outcome assessments to include the full range of possible cognitive and motor outcomes? Otherwise the model, while including the full range of severity (mild to severe), gets reduced to only include a segment of those who make significant recovery in the first 6 months. 6) Similarly, ability to comprehend (or lack of ability) might be an important cognitive outcome. Why are the severe cognitive impairment outcomes not included? 7) There are a very large number of input variables and they are not very well defined. e.g. Neurosurgical procedures - what will be included in the model? That a procedure was done? Or type of procedure? 8) Is the discussion section purposely left out?

REVIEWER	Catroppa, Cathy Murdoch Children's Research Institute
REVIEW RETURNED	23-Feb-2022

GENERAL COMMENTS	Thank-you for the opportunity to review this protocol paper interested in personalised prognosis for children with traumatic brain injury (TBI). The study will be of interest to those working with children post-TBI, but a few comments and queries have been raised below: Introduction: 1. There is much research which has looked at predictors of outcome following childhood TBI, with areas considered including injury (including imaging/advanced imaging), pre-morbid status, family, and environmental variables. These are also the variables included in the current protocol therefore, - Will this study add to current data? If so, mention and references of previous studies will be useful. If the authors feel that, while the variables are similar to other studies, they are being considered in an alternative way, then this could be made clearer. - If the machine learning component is novel, perhaps this can be extended and highlighted further? Method: 1. How were the medical centre/hospitals chosen? 2. How were the schools selected? 3. Were comorbidities an exclusion for the clinical or control groups? 4. Was an IQ below 70 an exclusion criterion? Sample size estimation: 1. I found the description of the sample size estimation confusing with regard to estimating 150 children (10x15) and then the calculation to reach 210 children. More information/explanation will be useful. Measures: 1. Was the mental health of the parents considered? MRI: 1. Are there any limitations in using one MRI scan result only at the one-month time point post-injury? Procedure: 1. What gift will be given to participants? 2. Have the authors considered collection of any qualitative data with regard to functional outcomes? 3. Provide more information on the Dutch Pupil Monitoring System. 4. Do families get a report at the conclusion of the study?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Sudhin Shah, Weill Cornell Medical College Comments to the Author:
The inability to accurately prognosticate the myriad outcomes following heterogeneous TBI certainly warrants large data-driven studies such as the one proposed. There are, however, several concerns about this current protocol.

We thank dr. Sudhin Shah for the comments on our study protocol. Below we respond to the comments raised. In addition, we would like to point out that unfortunately we were not able to address all comments of dr. Sudhin Shah in the manuscript, since this would require a discussion section, which is not part of the BMJ Open protocol manuscripts (see Reviewer 1, #8). Nevertheless, we hope that our response displayed below satisfactory answers the reviewer's questions.

Reviewer 1, #1

While the MRI based analyses are to be explored for their additive contributions, is it the expectation that if found to be useful, it would translate to clinical practice? i.e. is access to MRI facilities and the very specific expertise needed for analyses scalable?

Our primary aim is to develop a prediction model based on premorbid functioning and clinical parameters, which is readily available in a clinical setting and therefore has strong feasibility for clinical implementation. Yet, our secondary aim is to investigate the added predictive value of advanced MRI for outcome in children with TBI. We acknowledge that a prediction model based on advanced MRI parameters will not be readily implementable in clinical practice.

Although 3T MRI machines are typically available in the great majority of hospitals, availability may be much lower elsewhere, in developing countries for example. However, we see an important role for clinical research to also look beyond direct clinical implementability. When the state of the literature would indicate that MRI-based prediction models have important added value for prognosis, this information may be used to adapt routing of follow-up for children with TBI to hospitals equipped with the necessary resources. In the more distant future, the accessibility of MRI scanners may greatly improve thanks to the development of a mobile MRI scanner, which has recently been approved by the FDA.

We are also aware that some measurements are sensitive to individual scanner characteristics, limiting scalability of a MRI-based prediction model across MRI sites. However, recent developments show promising potential to correct for local scanner effects on MRI measurements. For example, Karayumak et al. (2019) has explored the use of a voxelwise algorithm designed to harmonize diffusion data across MRI sites to counteract the previous issues related to MRI specific scanners and DTI acquisition.

Karayumak, S. C., Bouix, S., Ning, L., James, A., Crow, T., Shenton, M., …. & Rathi, Y. (2019). Retrospective harmonization of multi-site diffusion MRI data acquired with different acquisition parameters. *Neuroimage, 184,* 180-200. https://doi.org/10.1016/j.neuroimage.2018.08.073

Going mobile: FDA clears world's first bedside MRI scanner-on-wheels | Fierce Biotech

Reviewer 1, #2

How will the machine learning approaches translate to real-life implementation? i.e. if validated in this initial sample, how is this expected to be used when a future child is admitted and the clinicians need to predict that child's function?

Machine learning models can be implemented in online environments that are easily accessible for clinicians. Such interactive models allow a user to input the values for the relevant predictors and subsequently provide the predicted outcome. For examples of interactive models, see Shiny - Gallery (rstudio.com)) and https://www.rstudio.com/blog/using-shiny-in-healthcare/.

Recently, machine learning models also have been implemented in electronic patient record environments (for a review see Si et al., 2021). In such cases, the model automatically extracts the relevant information from the patient's health record and automatically predicts outcome. Even automated actions can be coupled to certain predicted outcome. Recently during the COVID-19 pandemic, Epic (an electronic health record application) implemented a machine learning model that screens patients' health records and alerts doctors automatically before patients need an Intensive Care Unit admission or other care intervention (Schwab et al., 2021 also see Artificial Intelligence Triggers Fast, Lifesaving Care for COVID-19 Patients (epic.com)).

Schwab, P., Mehrjou, A., Parbhoo, S., Celi, L. A., Hetzel, J., Hofer, M., ... & Bauer, S. (2021). Real-time prediction of COVID-19 related mortality using electronic health records. *Nature communications, 12(1),* 1-16. https://doi.org/10.1038/s41467-020-20816-7

Si, Y., Du, J., Li, Z., Jiang, X., Miller, T., Wang, F., ... & Roberts, K. (2021). Deep representation learning of patient data from Electronic Health Records (EHR): A systematic review. *Journal of Biomedical Informatics, 115,* 103671. https://doi.org/10.1016/j.jbi.2020.103671

Reviewer 1, #3
What is the justification for excluding those below 4?

The age criterion was chosen to target children as from primary school age. In the Netherlands, children start primary school at the age of four. We have specified this in the manuscript as follows.

Methods and analysis, Study Population, page 8
Thereby, seeking to recruit a representative sample of children with TBI **from primary school onwards**.

Reviewer 1, #4
What is mild TBI criteria? concussions or complicated mild?

In this study, we do not use TBI severity as a selection criterion. All children who are clinically diagnosed with TBI by the attending paediatrician or paediatric neurologist in the emergency department are eligible for participation. National guidelines for clinical care for paediatric mild TBI are enforced in the emergency department, which define mild TBI in terms of a score 13-15 on the first assessed Glasgow Coma Scale, loss of consciousness ≤30 minutes, and posttraumatic amnesia <24 hours (Nederlandse Vereniging voor Kindergeneeskunde [Dutch Association of Pediatrics] 2010; Lumba-Brown et al., 2018). Because we do not use a research diagnosis for TBI, this may result in a more heterogeneous sample of children with TBI due to potential practice variation in adherence to diagnostic criteria. The used inclusion protocol including all children clinically diagnosed with TBI will enhance external validity of our study, with findings likely to better generalize to clinical practice.

Nonetheless, we will use TBI severity as a predictor, where we will follow the TBI classification system based on Tsao (2020), with a clinical diagnoses of mild, moderate or severe TBI according to the Glasgow Coma Scale score (13–15, 9–12 and 3–8, respectively), loss of consciousness duration (<1 hour, 1-24 hours and >24 hours, respectively) and post-traumatic amnesia duration (<1 day, 1–7 days and >7 days, respectively). We have specified this in the manuscript as follows.

Methods and analysis, Study Population, page 8-9
We will use a clinical diagnosis of TBI instead of a research diagnosis of TBI for inclusion in the TBI group. Although this may lead to a more heterogeneous study sample of children with TBI due to practice variation in the adherence to national guidelines stipulating criteria for assessment and treatment of mild TBI (Nederlandse Vereniging voor Kindergeneeskunde [Dutch Association of Paediatrics, 2010), this will also result in a study sample that better represents the clinical population of children with TBI. Exclusion criteria relating to very poor motor or cognitive outcome (exclusion criteria 2 and 3) will be registered as an outcome, to investigate the potential bias that may be introduced by lack of outcome assessment in a specific subsample of children with very poor outcome.
…
Inclusion criterion
5. ~~Hospital admission for~~ **A clinical diagnosis of** mild to severe TBI **according to a pediatrician or pediatric neurologist.**

Lumba-Brown, A., Yeates, K. O., Sarmiento, K., Breiding, M. J., Haegerich, T. M., Gioia, G. A., ... & Timmons, S. D. (2018). Diagnosis and management of mild traumatic brain injury in children: a systematic review. *JAMA pediatrics*, *172*(11), e182847-e182847. https://doi:10.1001/jamapediatrics.2018.2847

Nederlandse Vereniging voor Kindergeneeskunde [Dutch Association of Paediatrics]. (2010). Richtlijn opvang van patiënten met licht traumatisch hoofd/hersenletsel [Guideline on acute care for patients with mild traumatic head/brain injury] Link

Tsao, J. W. (Ed.). (2020). Traumatic brain injury: A clinician's guide to diagnosis, management, and rehabilitation. *Springer Nature.* (eBook) https://doi.org/10.1007/978-3-030-22436-3 p.2

Reviewer 1, #5
Exclusion criteria - why is the study excluding anybody who hasn't recovered to the point that they can participate in outcome assessments? i.e. if the purpose is to predict heterogeneous outcome, why exclude those who might still be recovering but might be in a disorder of consciousness or post-traumatic confusional state etc at the time of the assessments? Why not expand the outcome assessments to include the full range of possible cognitive and motor outcomes? Otherwise the model, while including the full range of severity (mild to severe), gets reduced to only include a segment of those who make significant recovery in the first 6 months.

We thank the reviewer for this valuable comment and understand the concern for a possible bias in the prediction model. According to our exclusion criteria, we will not exclude any participants upfront. However, if parents indicate, or test leaders notice, that a given participant lacks the ability to follow instructions, or has a severe motor disability that interferes with outcome assessment, those participants will be excluded from the outcome assessment. As suggested by the reviewer, this will be registered as an outcome, and we will investigate whether this influences the validity of our model. To elaborate, we will register whether outcome assessment was stopped due to severe disabilities and in case this number will be substantial we could also explore the development of a prediction model for very poor motor and/or cognitive outcome (precluding outcome assessment per protocol). We refer the reviewer to the previous comment (Reviewer 1, #4) on how this is now specified in the manuscript. Based on available outcome literature, we expect more than 95% of the children in our study cohort able to complete outcome assessment (Keenan & Bratton, 2006).

Keenan, H. T., & Bratton, S. L. (2006). Epidemiology and outcomes of pediatric traumatic brain injury. *Developmental neuroscience, 28(4-5),* 256-263. https://doi.org/10.1159/000094152

Reviewer 1, #6
Similarly, ability to comprehend (or lack of ability) might be an important cognitive outcome. Why are the severe cognitive impairment outcomes not included?

We refer the reviewer to our response to the previous comment (Reviewer 1, #5).

Reviewer 1, #7
There are a very large number of input variables and they are not very well defined. e.g. Neurosurgical procedures - what will be included in the model? That a procedure was done? Or type of procedure?

It is correct that our study proposes a wide range of input variables to predict outcome in children with TBI. This was purposely set up in order to optimize the potential of this study to capture the complex constellation of determinants that predict outcome. In order to identify the most relevant set of predictors (which is likely to differ per outcome variable) we will use a data driven approach, also utilizing the promising potential of machine learning models. Therefore, the selection of input variables is not determined upfront. Nonetheless, in Appendix 1 we now list a comprehensive yet concise (considering the readability) set of variables that we register in the hospital setting. In response to the reviewer's point regarding the registration of neurosurgical procedures, below we list how we distinguish between the following type of neurosurgical procedures (and how this is registered).

Neurosurgical assessment:

Not assessed
Assessed;
Neurosurgical intervention (if assessed)
- *No neurosurgical intervention necessary*
- *Neurosurgical procedure performed;*
  - *Relieve hematoma (craniotomy or Burr hole)*
  - *Decompressive craniectomy*
  - *Placing pressure gauge*
  - *Placing external CSF drain*
  - *Placement of ventriculoperitoneal drain (VPD)*
  - *Lumboperitoneal drain placement*
  - *III-ventriculostomy*
  - *Surgical treatment for persistent CSF leakage*
  - *Cranioplasty*

The following text has been added to the manuscript.

Methods and analysis, Protocol, page 10
Table 2 provides an overview of all study measures that will be recorded during hospital admission. **See Appendix 1 for a full listing of clinical measures that will be collected.**
Reviewer 1, #8
Is the discussion section purposely left out?

In accordance with BMJ Author Guidelines for protocol manuscripts we purposely left out the discussion section.
BMJ Author guidelines for protocols

Reviewer: 2

Dr. Cathy Catroppa, Murdoch Children's Research Institute Comments to the Author:
Thank-you for the opportunity to review this protocol paper interested in personalised prognosis for children with traumatic brain injury (TBI). The study will be of interest to those working with children post-TBI, but a few comments and queries have been raised below:

We thank dr. Cathy Catroppa for the valuable comments and queries on our study protocol.

Reviewer 2, #1
Introduction:
There is much research which has looked at predictors of outcome following childhood TBI, with areas considered including injury (including imaging/advanced imaging), pre-morbid status, family, and environmental variables. These are also the variables included in the current protocol therefore, Will this study add to current data? If so, mention and references of previous studies will be useful. If the authors feel that, while the variables are similar to other studies, they are being considered in an alternative way, then this could be made clearer.

The existing body of literature on predictors of outcome following childhood TBI has indeed provided the fundament for our current study protocol. With the PEPR study, we aim to build upon the existing literature by developing more personalized prediction models of outcome. To our best knowledge, this study is unique in the wide range of predictors assessed, covering a range of relevant domains. In addition, the existing literature on paediatric TBI, has rarely used machine learning approaches to build outcome prediction models. Machine learning has promising potential to improve the representation of complex relations (and interactions) between predictors and outcome. Lastly, although there is a considerable body of literature on the use of DTI (Dennis et al., 2021a; Dennis et al., 2021b), literature on the predictive value of structural and functional network parameters and neurometabolites is scarce, let alone when combined altogether. We argue that our study is also unique in the approach to investigate the added predictive value of a multimodal set of advances MRI predictors (i.e. combining, diffusion tensor imaging, resting state functional Magnetic Resonance Imaging and Magnetic Resonance Spectroscopy). To clarify the expected additional value of our findings we revised the manuscript as follows.

Introduction, page 5-6
Third, advanced **multimodal** magnetic resonance imaging (MRI) (**i.e. targeting brain volume, white matter integrity, structural and functional connectivity and neurometabolites**) has **not** ~~scarcely~~ been integrated in the existing prediction models, while ~~novel measures of structural and functional brain integrity and connectivity~~ **each of these MRI techniques has shown** ~~individual techniques show have shown~~ promising prognostic potential **when studied in isolation**.
…
Existing models have traditionally been developed using conventional statistical methods (e.g. logistic and linear regression) which may not harvest the full predictive potential of rich data sources in complex real-life outcome prediction.[25] Machine learning offers alternative models with high flexibility (e.g. decision trees, support vector machines), allowing more accurate data modelling.[25] Indeed recent application of machine learning in the prediction of global outcome after pediatric TBI has shown to improve the accuracy of prediction as compared to conventional statistical models.[26,27] Yet to date, the value of machine learning for the development of prognosis on more fine-grained yet crucial outcome domains in multifactorial disease conditions such as TBI, remains largely unexplored.[28,29]

a) Dennis, E. L., Caeyenberghs, K., Asarnow, R. F., Babikian, T., Bartnik-Olson, B., Bigler, E. D., ... & Wilde, E. A. (2021). Challenges and opportunities for neuroimaging in young patients with traumatic brain injury: a coordinated effort towards advancing discovery from the ENIGMA pediatric moderate/severe TBI group. *Brain imaging and behavior, 15*(2), 555-575. https://doi.org/10.1007/s11682-020-00363-x

b) Dennis, E. L., Caeyenberghs, K., Hoskinson, K. R., Merkley, T. L., Suskauer, S. J., Asarnow, R. F., ... & Wilde, E. A. (2021). White matter disruption in pediatric traumatic brain injury: results from

enigma pediatric moderate to severe traumatic brain injury. *Neurology*, *97*(3), e298-e309. https://doi.org/10.1212/WNL.0000000000012222

Reviewer 2, #2
If the machine learning component is novel, perhaps this can be extended and highlighted further?

We would like to refer to our response to the previous comment (Reviewer 2, #1).

Reviewer 2, #3
Method:
How were the medical centre/hospitals chosen?

All hospitals with an emergency department that provide medical care to children with mild to severe TBI are eligible for participation in this study. We have contacted trauma-level 1 centers, and general hospitals in the geographical area of our own hospital in order to build up a research network of hospitals participating in the PEPR study. This information was added to the manuscript as follows.

Methods and analysis, Study population, page 7-8
This study will prospectively recruit a multicenter cohort of children diagnosed with mild to severe TBI from a research network of ~~trauma-level-1 Dutch University Medical Centers and general~~ hospitals. **Trauma-level 1 Dutch University Medical Centers and general hospitals in the geographical area of the Amsterdam University Medical Center (initiator site) qualify as a participating center**.

Reviewer 2, #4
How were the schools selected?

Schools in the geographic area of the participating centres were contacted. However, the COVID-19 pandemic had a huge impact on the Dutch school system and led to an extreme high burden on school staff. Therefore, recruitment of the control group was subsequently mainly focused via contacting families through out-of-school care facilities, sports clubs and through existing collaborations with healthcare institutions. We added this information in the manuscript.

Methods and analysis, Study population, page 8
Demographically matched neurologically healthy children will be recruited, mainly via schools, **out-of-school care facilities, sports clubs and through existing collaborations with healthcare institutions in the geographic area of the participating centers.**

Reviewer 2, #5
Were comorbidities an exclusion for the clinical or control groups?

Per exclusion criteria in the Methods and Analysis section (Table 1, page 8), children with comorbidities are not excluded from the study. Thereby, we hope to recruit a representative sample of children with TBI that reflects the heterogeneity of the actual TBI population.

Reviewer 2, #6
Was an IQ below 70 an exclusion criterion?

The exclusion criteria are in the Methods and Analysis section (Table 1, page 8), IQ below 70 was not used as an exclusion criteria. Nonetheless, if participants were not able to comprehend test instructions at time of assessment, the test leader will not execute or stop the outcome assessment. Also, please see our response to Reviewer 1, #5.

Reviewer 2, #7
Sample size estimation:
I found the description of the sample size estimation confusing with regard to estimating 150 children (10x15) and then the calculation to reach 210 children. More information/explanation will be useful.

*We aim to develop an advanced, yet clinically relevant prognostic model for outcome in children aged 4-18 years with TBI. According to EMGO+ guidelines (EMGO+, 2015), 10-15 observations (i.e. participants) are required per predictor in the model (15 was chosen for a liberal calculation of the required sample size). Model complexity was set to a maximum of 10 predictors, in turn defining the minimum required sample size at (10 * 15 =) 150 children. According to Dutch medical ethical guidelines for scientific research, MRI will only be collected in children aged ≥8 years. Consequently, the model that will investigate the added predictive value of advanced MRI for outcome in children with TBI is on those children in the 8-18 year age range. This sets the minimum sample size per age year to (150 / (18 years – 8 years) = 15 children per age year. In turn, that results in a minimum total sample size of (15 * (18 years – 4 years) =) 210 children with TBI for the study sample of children aged 4-18 years.*

EMGO+. (2015). Progostic & Diagnostic Tests. Quality Handbook v 2.0.

Methods and analysis, Sample size estimation, page 9
Consequently, 10-15 observations are required per predictor in the model (15 was chosen for a liberal calculation of the required sample size). The minimum required sample size was calculated for an advanced, yet clinically relevant and implementable prognostic model. Hence, the model complexity was set to a maximum of 10 predictors, in turn defining the minimum required sample ~~size of the smallest subsample~~ at (10 * 15 =) 150 children. **Considering Dutch medical ethical guidelines for clinical research, children as from 8 years of age are eligible for MRI scanning in research. Therefore, we set the minimum required sample size (150 children) as the target sample size for the MRI subsample of children** (~~i.e. the subsample of TBI children~~ 8-18 years old). **Considering that we also aim to recruit an age-balanced TBI sample, we calculated the target sample of children with TBI per age year in the MRI subsample aged 8-18 years old (150 / (18 years – 8 years) = 15 children per age year) and applied this to the age range of the whole study sample aged 4-18 years old** ~~that are eligible for MRI assessment, also see 'Protocol: MRI assessment'). Thus, the sample size per age year is equal to (150 / (18 years – 8 years) = 15 children that results in a minimum sample size of~~ (15 * (18 years – 4 years) = **210 children) to arrive at the target sample size in the whole study sample of N =** 210 ~~children in the whole TBI group~~.

Reviewer 2, #8
Measures:
Was the mental health of the parents considered?

One questionnaire filled in by the parent of the participant assesses family functioning and will be used as a predictor of outcome. This questionnaire contains 28 statements and provides insight into

their household, parenting style, social relationships, the presence of a support system and on the parents' own childhood. Direct questions on mental health are not included, however we are confident that this measurement will be sensitive to the mental health situation of parents.

Methods and analysis, Protocol, page 10
Questions on premorbid functioning (assessing family[33] and behavioral[34] functioning) will be collected at the time of hospital admission to limit the contamination of assessment of premorbid functioning with potential consequences of TBI. The chosen questionnaires will lend beneficial insights in possible mediating factors of outcome, **assessing the presence of a social support system** and allows adjusting for SES, which is known to be important for TBI recovery.[35]

Veerman, J.W., Janssen, J., Kroes, G., De Meyer R.E., Nguyen L. & Vermulst A.A. (2012) 'Vragenlijst Gezinsfunctioneren volgens Ouders (VGFO). Handleiding.' *Praktikon*

Reviewer 2, #9
MRI:
Are there any limitations in using one MRI scan result only at the one-month time point post-injury?

Indeed, there are practical and technical reasons to assume that repeated scanning may provide valuable additional information. However, we selected one MRI assessment since this would greatly improve clinical implementability as compared to repeated MRI testing. In the case of more severe TBI, the child may not be fit for MRI scanning at the planned time point. Therefore, we have built in a time window of two weeks around the planned time point for MRI assessment, providing some flexibility to collect the relevant MRI data for children with more severe TBI.

Methods and analysis, Protocol, page 11
For eligible participants, one MRI session will be planned at one-month post-injury **with a two-week time window** at the Spinoza Centre for Neuroimaging, situated at the campus of the Amsterdam UMC. The chosen time point **of the** MRI assessment reflects a compromise between early measurement and the potentially confounding influence of brain oedema on advanced neuroimaging during the acute phase. Moreover, this ~~timing~~ time window aligns with routine follow-up of children after hospital admission for TBI according to the Dutch clinical guideline **and enables MRI assessment of children with more severe injuries.**

Reviewer 2, #10
Procedure:
What gift will be given to participants?

Next to the compensation of all travel expenses for all study procedures children will be thanked for their participation by offering them a small present (worth around €5,- for children aged 4-11 years and worth around €10,- for children aged 12 years or older). Participants can choose out of a small selection of age-appropriate presents (e.g. colouring books, wooden brain games, memory cards, sports gloves, ping pong set). We added this information to the manuscript.

Methods and analysis, Functional outcome assessment, page 14
To thank children for participation they will be given a small present after being debriefed and travel expenses will be reimbursed. **Participants can choose out of a small selection of age-appropriate presents (worth around €5,- for children aged 4-11 years e.g. colouring books, wooden brain games and worth around €10,- for children aged > 12 years e.g. sports attributes, card games).**

Reviewer 2, #11
Have the authors considered collection of any qualitative data with regard to functional outcomes?

This study is mainly focused on objective measures and will not collect any qualitative data with regard to functional outcome.

Reviewer 2, #12
Provide more information on the Dutch Pupil Monitoring System.

Thank you for addressing the need for further clarification. We added the following information to the manuscript.

Methods and analysis, Functional outcome assessment, page 15-16
School functioning will be assessed in the subsample of children attending primary school. Dutch Pupil Monitoring System[52] results will be requested through primary school teachers and include information prior to the injury as well as six months-post injury. **The Dutch Pupil Monitoring System developed by the National Institute of Educational Measurement in the Netherlands is to obtain reliable data systematically on pupil learning progress during their entire primary school career (Vlug, 1997). Test packages are developed for all 6 age groups between 6 and 12-years-old and allow seamless charting of academic development across these ages. We will assess packages developed for arithmetic's, spelling and technical reading (Glas & Geerlings, 2009).**

Glas, C. A. W., & Geerlings, H. (2009). Psychometric aspects of pupil monitoring systems. Studies in Educational Evaluation, 35(2–3), 83–88. https://doi.org/10.1016/j.stueduc.2009.10.006

Vlug, K. F. M. (1997). Because every pupil counts: The success of the pupil monitoring system in the Netherlands. Education and Information Technologies, 2(4), 287–306. https://doi.org/10.1023/A:1018629701040

Reviewer 2, #13
Do families get a report at the conclusion of the study?

For families interested in the conclusion of the study a concise report will be provided. In addition, on request families can retrieve a report with individual outcomes for measures with readily available normative data (e.g. subtest of the Wechsler Intelligence Scale). We added this information to the manuscript.

Ethics and dissemination, page 19
Upon completion of all study measures for all participants, **we will provide families interested in the results of the study with a concise report. In addition, on request, families can retrieve a report with individual outcomes for measures with readily available normative data (e.g. subtest of the Wechsler Intelligence Scale).**

The research data including a manuscript will be published in international peer-reviewed journals, preferably open-access.

Reviewer: 1
Competing interests of Reviewer: N/A

Reviewer: 2
Competing interests of Reviewer: n.a.

Appendix 1

1. Emergency care
    - 1.1. Arrival
        1. Date and time
        2. Trauma Mechanism
        3. Suspicion of abuse
        4. Medication at the emergency department (including helicopter/ambulance)
    - 1.2. Advanced Paediatric Life Support Protocol
        1. Awareness Level (Lowest GCS Score)
        2. Loss of consciousness
        3. Behaviour
        4. Anterograde Post-Traumatic Amnesia
    - 1.3. Blood gas analysis and physiological parameters
        1. Blood gas analysis
        2. Physiological parameters
    - 1.4. Indications CT
        1. Indications CT 2-5 years
        2. CT indications from 6 years
    - 1.5. Abbreviated neurological examination
        1. Pupils
        2. Eye movements
        3. Face
        4. Motor skills
        5. Reflexes

2. Intensive Care Unit (if applicable)
    - 2.1. Admission
        1. Admitted to the intensive care unit
        2. Date
    - 2.2. Measures
        1. Mechanical ventilation
        2. Trachea cannula
        3. Device data
        4. Glasgow Coma Scale score on admission
        5. Pupil size on admission
        6. Pupil reaction on admission
        7. Intracranial Pressure
        8. Medication
    - 2.3. Discharge
        1. Date
        2. Discharge destination
        3. Discharge Condition

3. Nursing ward (if applicable)
    - 3.1. Admission
        3. Admitted to the Nursing Ward
        4. Date
    - 3.2. Measures
        1. Mechanical ventilation
        2. Trachea cannula
        3. Device data
        4. Glasgow Coma Scale score on admission

5. Pupil size on admission
6. Pupil reaction on admission
7. Intracranial Pressure
8. Medication

3.3. Discharge
1. Date
2. Discharge destination
3. Discharge Condition

4. Radiology (if present)

4.1. CT scan
1. CT scan performed
2. Abnormalities on CT scan
3. Rotterdam CT score

4.2. MRI scan
1. MRI scan performed
2. Abnormalities on MRI scan

5. Neurosurgery

5.1. Neurosurgical Assessment

5.2. Neurosurgical Intervention (if assessed)

VERSION 2 – REVIEW

REVIEWER	Shah, Sudhin Weill Cornell Medical College
REVIEW RETURNED	18-Mar-2022

GENERAL COMMENTS	Thank you for responding to the previous review. I have no further comments.