Journal of Neuro-Oncology
https://doi.org/10.1007/s11060-020-03594-6

CLINICAL STUDY

Development of a clinical scale for assessment of patients with diffuse intrinsic pontine glioma (DIPG) receiving experimental therapy: the PONScore

Milo Alexander Hollingworth[1] · **Stergios Zacharoulis**[2]

Received: 12 May 2020 / Accepted: 8 August 2020

Abstract

Purpose Monitoring neurological side-effects in experimental therapy for diffuse intrinsic pontine glioma (DIPG) can be challenging. We aimed to develop a neurological scale that could be used by non-specialists to quantify neurological changes during experimental treatment of DIPG.

Methods We developed the Pontine Observational Neurological Score (PONScore) to measure signs and symptoms of DIPG by adapting validated assessment scales of neurological signs and symptoms in children. We developed a prototype score, taught it to paediatric intensive care nursing staff, who used the Score to assess children receiving awake pontine infusion of chemotherapy for treatment of DIPG. We used their feedback to develop the PONScore. Points are allocated for headache, ophthalmoplegia, facial and tongue weakness, dysarthria, paraesthesia, limb weakness and dysmetria with increasing scores reflecting increasing disability. The PONScore was administered every hour during awake pontine infusion. Correlation and agreement calculations between nursing staff, as non-specialists, and a specialist rater were performed in 30 infusions in 6 children (aged 8–11). Changes in PONScore versus volume of infusion are described in a further 55 infusions in 8 children (aged 3–11).

Results The PONScore demonstrated excellent intra-rater reliability with an intra-class co-efficient of 0.98 (95% CI 0.97–0.99; p-value < 0.001) between a specialist and non-specialist raters with strong correlation between scores and a Spearman correlation coefficient of 0.985 (p < 0.001). PONScores increased from 3.3 to 5.7 (p-value < 0.001) during infusion reflecting accumulation of neurological signs and symptoms during infusion.

Conclusions We describe a novel neurological scale that can be used by non-specialists to describe acute neurological changes in children receiving experimental therapy for DIPG. Prospective validation as part of a clinical trial is required.

Keywords DIPG · Midline diffuse glioma · Convection enhanced delivery · Neurological assessment · Clinical scales

Introduction

Diffuse intrinsic pontine glioma (DIPG) is a lethal disease of middle childhood, characterised by progressive neurological disability and worsening quality of life [1]. The tumour, thought to arise from precursor oligodendrocytes in the brain stem [2], injures adjacent structures leading to complex combinations of limb weakness, ataxia and cranial nerve palsies [1]. Trials so far have failed to show survival benefit from any treatment beyond conventional radiotherapy [3]. As such, radical treatments are being developed including intra-arterial drug delivery [4], convection enhanced delivery (CED) [5–7] and CAR T-cell therapy [NCT04196413]. The increasing understanding of the molecular genetic basis of the disease has also led to a re-emergence of brain stem biopsy, which was once defunct owing to an unacceptably high morbidity rate [8]. Emerging capabilities have created a new optimism for DIPG [9] but how best to monitor and optimise these new treatment strategies to reduce side-effects is unknown.

✉ Milo Alexander Hollingworth
milo.hollingworth@nuh.nhs.uk

[1] Department of Neurosurgery, Queens Medical Centre, Nottingham University Hospital NHS Trust, Nottingham NG7 2UH, UK

[2] Paediatric Haemato-Oncology, CUMC/Herbert Irving Pavilion, New York, NY 10032, USA

Published online: 09 September 2020

There are several challenges to consider when monitoring children receiving experimental treatment for DIPG. DIPG is rare, with 200–400 new cases per year in the United States [10]. Experimental treatments are carried out in national and international referral centres, which leads to a pooling of expertise faraway from patients' local care providers. Patients may have to travel long distances for follow-up. The progression of DIPG and pontine injury from therapies with preferentially local effects may present with new or worsening neurological impairment giving rise to complex neurological phenotypes. These patterns of deterioration will be unique to each patient, which could obscure deleterious effects of treatment and require specialist expertise to assess. Potential complications may therefore present in hospitals without relevant expertise and they may require guidance about how best to evaluate these patients.

Hence, it would be beneficial to establish a common method of how to assess these patients so that safety issues arising from treatment can be identified and compared. But, conventional methods of neurological assessment, such as the Glasgow Coma Scale (GCS) or grading of limb weakness, may only change when irreversible damage has taken place or may not elicit the complexity of deficits caused by pontine injury. Assessment by a paediatric neurologist would represent an ideal alternative; however, such expertise is not always readily available nor can changes in neurology be easily quantified. A pragmatic solution would be to develop a method of quantifiable assessment that can be elicited by non-specialists.

We treated patients with DIPG using chemotherapy administered by CED on compassionate grounds [7]. CED is a method of continuous positive pressure intra-parenchymal infusion delivered by one or more indwelling catheters [11]. It can bypass the blood brain barrier, distributing drug through brain volumes 3–7 times the volume of infusion [12]. Based on evidence from preclinical studies, we infused carboplatin and sodium valproate, as monotherapies and combination therapies, on compassionate grounds as a prelude to a clinical trial [7, 13]. CED of drug into the pons is associated with neurological side-effects [5–7]. Souweidane et al. (2018) treated 37 patients with DIPG in their landmark Phase I study of 124I-8H9 radio-labeled monoclonal antibody delivered by CED under general anaesthesia [6]. Twenty-five percent of side-effects recorded were consistent with neurological signs and symptoms of brain stem dysfunction e.g. limb weakness, cerebellar dysfunction and ophthalmoplegia. It is likely that this reflects the local effects of drug delivery. However, these acute changes could be due to the mechanical effect of infusion or the pharmacology of the drug. In the longer term, delayed neurological deterioration after pontine infusion could also be due to disease progression. Differentiating between these processes is challenging, particularly when our understanding of radiological changes after direct intra-parenchymal drug infusion are still developing. Dissecting the clinical impact of the drug, the infusion and the disease is fundamental to understanding CED as an experimental treatment. We developed the Pontine Observational Neurology Score (PONScore) to allow non-specialists to perform hourly quantifiable assessment of neurological signs and symptoms during infusion. We describe how the PONScore was developed and implemented and how it could be used to monitor other treatments for DIPG.

Methods

Ethical approval

Treatment of patients using CED on compassionate grounds was approved by the institutional ethics committee and was compliant with the 1964 Helsinki declaration and its later amendments. Implantation of the drug delivery system was approved by the Medicines and Healthcare Regulatory Authority, United Kingdom. Parents were consented for the experimental nature of the treatment and the use of their child's information for the development of the treatment and for scientific publication. Data was acquired from routinely collected information as part of their clinical treatment.

Pontine infusion

The method of chronic, intermittent CED has previously been described [5, 7, 14, 15]. In short, patients with a radiological diagnosis of DIPG were fitted with a 4-catheter chronic, intermittent drug delivery system using two transfrontal and two trans-cerebellar catheters connected to a bone-anchored transcutaneous drug administration port via sub-galeal tubing (Renishaw Drug Delivery System; Renishaw Plc, Wooton-under-Edge, UK). This allowed repeated awake pontine infusions without the need for repeated surgery. Patients would sit in bed connected to the infusion set for the duration of the infusion connected by two metre long extension lines. Infusions were supervised in a Pediatric Intensive Care Unit (PICU) where they received continuous non-invasive cardiorespiratory monitoring and hourly neurological assessment. Based on prior experience [5, 14], infusions would be continued until the onset of neurological symptoms to maximise the volume of tumour treated. Each treatment would typically consist of two infusions over two separate days and were repeated at 3–8 weekly intervals depending on patient fitness.

Development of the PONScore

Our clinical experience of pontine infusion suggested that the most common side-effects during infusion were

headache, long tract signs, ataxia and cranial neuropathy (III–XII). We identified recently validated clinical scales that could be feasibly conducted by the bedside in a child connected to an infusion. A literature search was conducted using PubMed with the search heading: ((scale stroke) OR scale ataxia) OR scale cranial nerve) AND ((childhood) OR (pediatric)). Only journal articles published within the last 10 years studying humans were included. Articles were excluded if they did not include clinical assessment scales or did not assess long tract signs, ataxia or cranial neuropathy (III–XII). Scales were selected if they had been validated in a paediatric population and could be conducted at the bedside without equipment. Three hundred and eighty-two articles were identified. Eight articles included clinical assessment scales. Three scales were excluded because they were not validated in children. One scale was excluded because it required equipment. Four eligible scales included the pediatric National Institute of Health Stroke Scale (pedsNIHSS) [16], Scale for Assessment and Rating of Ataxia (SARA) [17, 18], Brief Ataxia Rating Scale (BARS) and International Co-operative Ataxia Rating Scale (ICARS) [18] (Fig. 1). The pedNIHSS is used as a measure of neurological injury due to stroke and assesses supratentorial and infratentorial neurological signs and symptoms. SARA, BARS, ICARS are measures of ataxia and cerebellar dysfunction. Scales were deconstructed into their individual items. Items that did not assess long tract signs, cranial neuropathy (III–XII) or ataxia were excluded. Items that could not be completed safely while attached to the infusion set were also excluded, i.e. measures of gait, stance and sitting balance. Items that tested the same neurological domain, i.e. ataxia and dysarthria, were selected for ease of use and expected sensitivity to change during infusion. The resulting items were amalgamated into a Prototype Score with addition of items for headache (Wong-Baker Pain Scale® [19]), facial nerve palsy (House-Brackmann Scale [20]) and a measure of lingual range of movement [21]. Inclusion of consciousness assessment was also included from the pedNIHSS to replace the need for simultaneous assessment of GCS. The prototype scale was administered by paediatric intensive care unit (PICU) nursing staff following bedside tutorials. PICU nurses had no prior specialist neurological training beyond their paediatric intensive care nursing accreditation. Following a trial period, nursing feedback was used to refine the Prototype into the PONScore. The PONScore was then taught to PICU nursing staff during weekly teaching and bedside tutorials. The PONScore was then performed hourly from the beginning to the end of the each infusion as part of the standard operating procedure of awake pontine infusion.

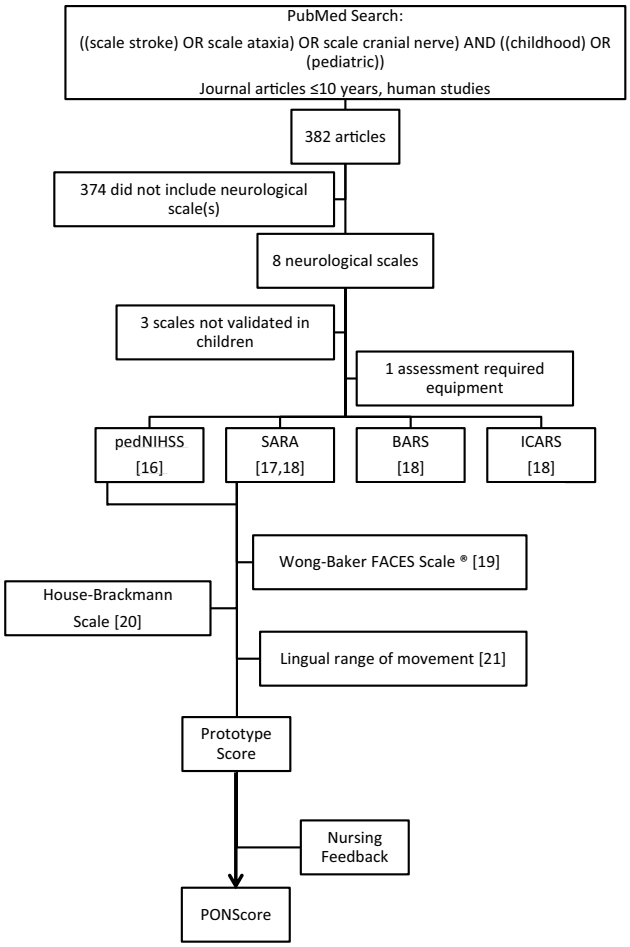

Fig. 1 The Pontine Observational Neurology Score PONScore was developed from existing clinical scales identified from a literature search. Addition of further items to measure headache (Wong-Baker FACES Scale [19]), lingual range of movement [21] and facial movement (House-Brackmann Scale [20]) were included in a prototype score. Based on nursing feedback of the prototype the PONScore was developed. Abbreviations: pediatric National Institute of Health Stroke Scale (pedsNIHSS) [16], Scale for Assessment and Rating of Ataxia (SARA) [17, 18], Brief Ataxia Rating Scale (BARS) and International Co-operative Ataxia Rating Scale (ICARS)

Validation and analysis of the PONScore

PONScore was analysed retrospectively using data collected as part of routine clinical care. Data was analysed by comparing assessments made by the attending nurse and doctor during 30 separate infusions (27 Sodium Valproate, 3 Carboplatin). These infusions were performed in 6 children with DIPG (aged 8–11). Patients were expected to develop neurological signs and symptoms during infusion. To ensure the PONScore performed equally well throughout infusion, nurse and doctor-recorded PONS scores were collected at two time points (at hour 0 and hour 6).

The overall PONScores recorded by nurse and doctor were compared for mean difference, correlation, inter-rater reliability and agreement. Mean PONScores were compared using Mann–Whitney U Test. Correlation was measured using Spearman correlation co-efficient. Inter-rater reliability was measured by calculating intra-class co-efficient (ICC) of the total score. Agreement of the total score was measured using Cronbach's alpha and Kendall's co-efficient. Scoring bias and the limits of agreement between the two raters' total scores were estimated using Bland and Altman methods [23]. Agreement between individual items was measured using Kappa values, Kappa ≤ 0.40, $0.40 > k < 0.75$ and $k \geq 0.75$ were used to define poor, moderate and excellent agreement respectively. Statistical analysis was performed using Statistical Package Social Sciences (Version 23, IBM). Statistical significance was defined at p-value < 0.05.

Results

Nursing feedback

Nurses trained in using the prototype score were asked to complete a paper survey based on their prior experience of performing awake pontine infusion. Using a Likert scale, 9/10 nurses either strongly agreed or agreed that the prototype assessment provided useful information during pontine infusion and 8/10 agreed or strongly agreed it could be easily performed with their existing workload. 6/10 nurses said assessment took 1–2 min and 9/10 of nurses felt confident or very confident using the assessment. Particular comments requested scoring criteria for mood changes, subtle limb weakness not otherwise recorded and sensory symptoms. They also requested specific instructions about how to identify dysarthria, restriction of eye and facial movements.

Based on this feedback, the prototype was adjusted into the final PONScore (Table 1). A point was included for distress during infusion. Additional criteria were included to assess drift < 10 cm in the arm or leg and occurrence of sensory symptoms. The item for assessment of eye movement was also adjusted to specify how restricted eye movements should be elicited. By trying to bury the limbus, extreme lateral gaze would be elicited and each eye would be inspected for the ability to fully adduct and abduct each eye [22]. Visible sclera between the limbus and the epicanthus was defined as restricted gaze. Words from the adult NIHSS ("*mama, baseball player, fifty-fifty, thank, huckleberry, caterpillar*") were also included to standardise assessment of dysarthria. The principle of observing facial symmetry at rest and during movement when using the House Brackmann Scale was kept; however, the score was broadly modified to identify bilateral facial nerve palsy and to give higher scores for upper facial weakness to reflect the increased risk of corneal injury (Table 2).

Agreement and reliability of the PONScore

Mean PONScore was 5.49 and 5.5 recorded by nurses and doctor respectively (p-value = 0.727). Overall, agreement and reliability of the PONScore was calculated from 60 paired examinations—two from each infusion. There was strong correlation (0.985; p-value = 0.01) between nurse and doctor scores. Scores recorded by doctors and nurses were identical in 45/60 recordings and were within one point in 57/60. ICC was 0.98 (95% CI 0.97–0.99; p-value < 0.001). Cronbach's alpha value was 0.98. Kendall's Co-efficient was 0.97. Bias in scoring, estimated using Bland and Altman methods [23], was 0.03 and there was no evidence of proportional bias (Fig. 2). Kappa values demonstrated excellent agreement ($k \geq 0.75$) for all items except consciousness, which was incalculable because it did not change during any infusion. ICC, Kappa, Cronbach's alpha values were calculated for hour 0 and hour 6 separately, which demonstrated similar results, suggesting the reliability and agreement of the PONScore was stable throughout infusion.

PONScores from a further 55 infusions were analysed to describe the change in PONScore during infusion. Eight children (3–11 years) with DIPG received a total of 55 infusions of combined carboplatin and sodium valproate. Between 3.0 mL and 5.6 mL were delivered over 5–11 hours as patient tolerance allowed. Median PONScore at the start of each infusion was 2 (range 0–16). Changes in PONScore from pre-infusion baseline reflected the accumulation of neurological signs and symptoms during infusion (Fig. 3). Mean PONScore increased during infusion from 3.3 to 5.7, which differed with statistical significance (p-value < 0.001; Fig. 3).

Discussion

We developed and implemented the PONScore for assessment of neurological change during pontine infusion of chemotherapy for treatment of DIPG. This experimental therapy was associated with local brain stem signs and symptoms: headache, cranial neuropathy (III–XII), limb weakness and cerebellar ataxia. We modified previously validated measures of paediatric neurological function to develop this new score, which was taught to non-specialist nursing staff with bedside and seminar teaching. PONScores recorded by nurses demonstrated high degrees of agreement and inter-rater reliability with a specialist assessor. This helped to quantify the accumulation of neurological signs and symptoms during pontine infusion, which we hope will

Table 1 Pontine observational neurology score: examination guide and mark scheme

Component and examiner instructions	Criteria (Score)
Headache	
Nurse instruction: Ask: "Do you have a headache? How much does it hurt?" Examiner points to Wong-Baker Faces® [20]	No headache (0) Hurts a little (1) Hurts little more (2) Hurts even more (3) Hurts a whole lot (4) Worst hurt (5)
Consciousness	
Nurse instruction: Observe the patient, which description is most appropriate?	Keenly alert (0) Rousable but alert with light stimulation (1) Rousable with strong stimulation (2) Reflex movements or unresponsive (3)
Eye movement	
Nurse instruction: Ask the patient "Look left and right" and/or elicit extreme lateral gaze. Check pupillary light response in total gaze paresis. Ignore nystagmus.	Buries the limbus in the epicanthi of the direction of gaze on both sides (0) Restricted gaze in one eye (1) Restricted gaze in both eyes (2) Total gaze paresis; where you are unable to elicit any eye movements: Symmetric pupils that are responsive to light (3) Pupils that are asymmetric or unresponsive to light (4)
Facial sensation	
Nurse instruction: "Do you have numbness or tingling? If so, is it painful? If so, is it severe?"	No numbness or tingling or pain in the body or limbs (0) Non- painful sensory change (1) Moderately painful sensory change (2) Severely painful sensory change (3)
Facial movements	
Nurse instruction: Note the face at rest and on movement noting the upper and lower face. Ask the patient to "Raise your eyebrows, scrunch up your eyes, blow out your cheeks, show me your teeth"	Symmetrical face at rest, which is symmetrical throughout all movements (0) Symmetrical at rest with obvious asymmetry during movement involving the lower face only (1) Symmetrical at rest with obvious asymmetry during movement involving the upper face (2) Asymmetrical at rest with obvious asymmetry during movement involving lower face only (3) Asymmetrical at rest with obvious asymmetry during movement involving the upper face (4) Barely perceptible or no movement on one side of the face (5) No movement on either side of the face (6)
Tongue movements	
Nurse instruction: Ask the patient "Stick your tongue out as far as you can"	Normal range of movement (0) Deviation with normal protrusion beyond lower lip margin (1) Unable to protrude tongue beyond lower lip margin (2)
Body sensation	
Nurse instruction: "Do you have numbness or tingling? If so, is it painful? If so, is it severe?"	No numbness or tingling or pain in the body or limbs (0) Non-painful sensory change (1) Moderately painful sensory change (2) Severely painful sensory change (3)

help us understand risk factors for persistent neurological side-effects. As a measure of neurological signs and symptoms arising from brain stem dysfunction, the PONScore may also be used to evaluate neurological effects of other therapies for DIPG, particularly those that have local effects within the pons and adjacent structures.

Table 1 (continued)

Component and examiner instructions	Criteria (Score)
Arm power	
Nurse instruction: The limb is placed in the appropriate position: extend the arms (palms up) 90 degrees. Drift is scored if the arm falls before 10 seconds. Score each arm separately.	No drift; limb holds 90 degrees for full 10 seconds without drift (0) Minor drift; limb holds but drifts down <10 cm or pronates before full 10 seconds; does not hit bed or other support (1) Drift; limb holds but drifts down >10 cm before full 10 seconds; does not hit bed or other support (3) Some effort against gravity; limb holds but drifts down to bed or other support, but has some effort against gravity (3) No effort against gravity; limb falls (4) No movement (5)
Leg power	
Nurse instruction: The limb is placed in the appropriate position: leg straight with heel raised 30 cm above the bed. Maintain for 5 seconds. Score each leg separately.	No drift; limb held with heel 30 cm off bed for 5 seconds (0) Minor drift; limb held above bed but drift < 10 cm or external rotation at hip (1) Drift; limb held with heel above bed but drift >10 cm; does not hit bed or other support (2) Some effort against gravity; limb holds 30 cm above bed but drifts down to bed or other support but has some effort against gravity (3) No effort against gravity; limb falls (4) No movement (5)
Limb co-ordination	
Nurse instruction: The finger-nose-finger test and heel shin test	Normal co-ordination (0) Abnormal in one limb (1) Abnormal in two limbs (2) Abnormal in three limbs (3) Abnormal in four limbs (4)
Speech	
Nurse instruction: Consider the patients speech since your last assessment and ask the patient to pronounce: *'huckleberry, mama, fifty-fifty, thanks, baseball, caterpillar'*	Normal pronunciation and normal conversational speech (0) Suggestion of speech disturbance in conversational speech only (1) On pronunciation of *'huckleberry, mama, fifty-fifty, thanks, baseball, caterpillar':* Suggestion of speech disturbance during pronunciation (2) Slurring of words (3) The words are easily understood but there is obvious slurring (4) Difficult to understand (5) Noises only (6) Aphasic (7)

The patient is examined in bed in the semi-recumbent position with the head of the bed at 45 degrees

There are important limitations to this study. Awake pontine infusion in DIPG is a treatment limited to a very small number of patients administered by an even smaller group of healthcare professionals. This limited our ability to validate the score in a large number of patients. However, each infusion can represent a unique neuro-oncological phenotype and so the estimates of reliability may be more robust than the small number of children studied would suggest. Nevertheless, our reliability estimates would be improved by testing in patients with and without DIPG and in children of different ages. Notably, patients were mostly above 5 years old and an abbreviated version for children under 5 years of age should be developed. Patients did not have severe disability during testing and so the score may not be reliable at higher levels of disability. Indeed, patients who received CED may not represent all children with DIPG. Patients receiving CED must be fit enough to undergo surgery and have a tumour amenable to coverage by the implanted delivery system. Therefore, the PONScore may not be as reliable in patients with cystic necrosis, haemorrhage, advanced or disseminated disease. Also, the extent to which changes in the PONScore correspond to clinically significant change is unknown. The impact of first language on assessment of dysarthria has not been explored and should be considered

Table 2 Kappa values for agreement of the PONScore by each individual item

Item	Kappa values (standard error)		
	Hour 0	Hour 6	Overall
Headache[a]	1 (0.000)	1(0.000)	1 (0.000)
Consciousness[b]	–	–	–
Eye movements	0.813 (0.095)	0.945 (0.049)	0.891 (0.052)
Facial sensation[a]	1 (0.000)	1 (0)	1 (0)
Facial movement	0.935 (0.089)	0.946 (0.048)	0.94 (0.379)
Tongue movement	1 (0.000)	0.884 (0.101)	0.942 (0.054)
Body sensation[a]	1 (0.000)	1 (0)	1 (0.000)
Arm power	0.932 (0.064)	0.771 (0.087)	0.851 (0.055)
Leg power	0.918 (0.076)	0.895 (0.072)	0.916 (0.039)
Limb ataxia	0.924 (0.074)	0.883 (0.081)	0.906 (0.052)
Speech	0.796 (0.106)	0.836 (0.083)	0.828 (0.065)

[a]Self-reported items

[b]Consciousness did not change during any infusion

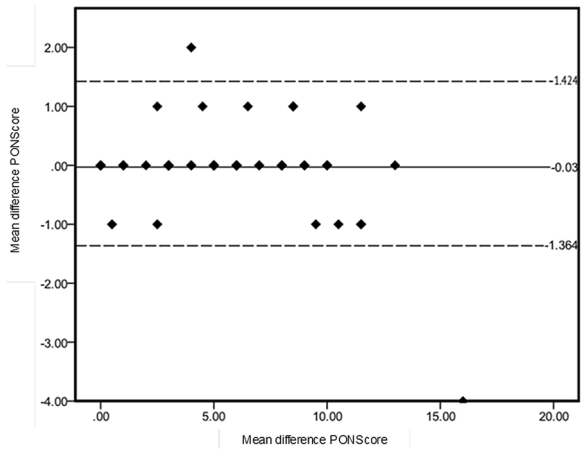

Fig. 2 A Bland Altman Plot [23] of study bias using the PONScore measured by a specialist doctor and non-specialist nurse in 30 awake pontine infusions

when conducting international trials. In addition, the estimates of reliability and agreement should be interpreted with caution given the retrospective nature of the data analysis, absence of blinding to the patients' infusion regime and the potential for doctors and nurses to confer during scoring. The reliability of the PONScore was only assessed during single infusions and hence the PONScore should be used alongside other measures of neurological function if it is used in long term follow-up.

It is clear that validation in a clinical trial is required. The PONScore was designed in the context of an experimental therapy in the early stages of its clinical development in a single international referral centre; this specific context requires special consideration. Boateng et al. (2018) describes a nine step, three-phase process for development and validation of new scales for health, social and behavioural research. The best practices they describe rely on larger numbers of participants and assume a broader base of pre-existing knowledge. As such, some recommendations they make were not feasible; for example, item development by a large expert panel was not possible due to the small number of clinicians involved in the development of the treatment. Nevertheless, conducting a thorough literature review, developing a scale from already validated measures of neurological function, which improves content validity, and involvement of end-users (i.e. nursing staff administering the score) all adhere to good practice when developing new scales [24]. We also argue that the PONScore is a practical solution to an important clinical problem. Experimental treatment for DIPG may cause unpredictable and complex neurological signs and symptoms, which could pose an immediate risk or lead to persistent disability. Recognising early signs of neurological compromise and understanding how acquired deficits relate to treatment is of paramount importance. As such, neurological examination in experimental treatment for DIPG must be performed regularly during treatment and elicit the complex neurological signs and symptoms that typify DIPG patients. Specialist paediatric neurological assessment may not always be possible, particularly if required frequently, nor can it be easily quantified or compared between different raters. We developed a score that quantified cardinal features of brain stem dysfunction in children with DIPG, which demonstrated high degrees of agreement and inter-rater reliability between a specialist and non-specialist rater who had received appropriate training. Although, further testing is required, we can be reassured that the score has validity because it is adapted from scales validated in children already in clinical use. Moreover, because the administering nurses were involved in the development of the scale, we believe the PONScore is practical. In the absence of a better alternative, the PONScore may be a useful tool to evaluate neurological change during experimental treatment for DIPG.

Limitation of existing measures of neurological function should also be considered when evaluating the potential applications of the PONScore in current clinical practice. In standard practice, bedside observation mandates pupilometry and GCS assessment [25], but this will be insensitive to the complexity of signs and symptoms in DIPG patients. The Response Assessment in Neuro-Oncology Group proposed the Neurological Assessment in Neuro-Oncology Scale to quantify gait, language, vision, limb strength, ataxia, sensation and behaviour in patients with brain tumours [26]. However, this scale would be impractical for patients receiving CED who are attached to an infusion set, it has not been validated in children with DIPG nor does it assess bulbar function. Indeed, the CTCAE, which is the standard measure

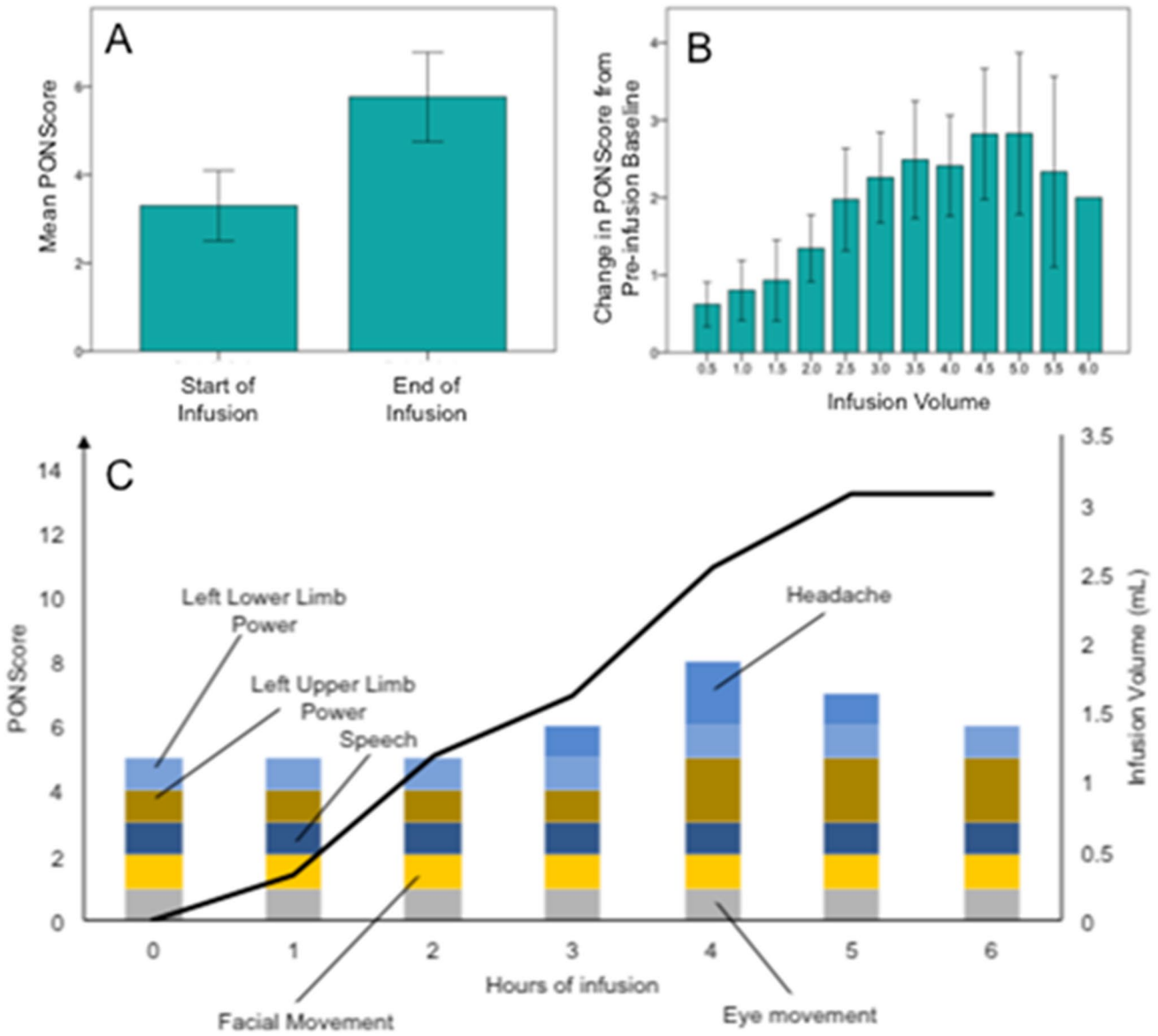

Fig. 3 Changes in neurological signs and symptoms measured using the Pontine Observational Neurology Score (PONScore) in 55 awake pontine infusions performed in eight children (3–11 years) with Diffuse Intrinsic Pontine Glioma (DIPG). **a** Mean PONScore increased from 3.29 to 5.65 from the start of infusion to the end of infusion, which reached statistical significance (p-value < 0.001). **b** Mean change in PONScore from pre-infusion baseline increases with increasing infusion volume. **c** A representative graph is shown of a patient with DIPG with a complex neurological phenotype at baseline who acquired new and worsening neurological signs and symptoms during infusion. Hourly PONScores are separated by each domain (coloured bars) versus volume of infusion (black line). At baseline the child, had mild left hemiparesis, right 6th nerve palsy and facial weakness. At hour 3 of infusion, the boy developed headache. At hour 4 of infusion, left upper limb weakness increased in severity. At hour 5, infusion was stopped. Headache resolved and no further neurological deterioration was observed. Error Bars: 95% confidence interval

of adverse events used in clinical trials, cannot be used to localise or identify acute neurological side-effects. In multi-catheter pontine infusion, where drug is infused into multiple parts of the brain stem simultaneously, it is important to identify which catheter is likely to be symptomatic so the infusion can be titrated appropriately. Moreover, the CTCAE quantifies neurological side-effects based on their associated disability, which may only be apparent once the patient has returned to their activities of daily living and cannot be used to evaluate rapid neurological change. Souweidane et al. [6] measured treatment related side-effects at seven days after infusion using CTCAE. This cannot delineate the effect of infusion versus drug toxicity nor can it inform how side-effects can be minimised acutely. The PONScore could provide a standardised neurological assessment that can follow the onset and recovery of neurological

side effects during and after treatment. This will form an important part of defining infusion-related and drug-related side-effects during CED treatment. Such definitions may allow us to develop rules about how to titrate infusion and also understand the risk factors that contribute to disabling side-effects. Importantly, the PONScore showed significant increases in neurological signs and symptoms during infusion. This emphasizes how performing pontine infusions awake enables monitoring of neurological function and, if performed meticulously, how this could maximise the safety of the procedure. Indeed, from analysis of 55 infusions it appears that changes in PONScore are less pronounced at higher infusion volumes. This counterintuitive observation requires further study; but it suggests that infusion-related side-effects are not simply dependent on infusion volume. Hence, in the future it may be possible to develop infusion regimes that are minimally symptomatic.

CED is not the only circumstance in which standardised neurological assessment is of benefit. When developing new treatments, it is necessary to demonstrate clinical benefit. A key priority for patients with brain tumours is maintaining neurological integrity [27, 28] and using a standardised neurological assessment such as the PONScore may serve as a short and long term outcome measure to evaluate other emerging therapies. The PONScore could be used to gain a deeper understanding of the type and severity of neurologic dysfunction associated with current treatments, particularly those with local effects. Comparison of neurologic changes after different radiotherapy regimes (hypofractionated versus conventional), brain biopsy techniques or steroid regimens could all be measured using the PONScore. Given that the PONScore is recorded based on self-reported symptoms and observed neurological signs it may be possible to validate the PONScore for televisual assessment, which could also reduce the burden of patient follow-up particularly in international trials.

In conclusion, we describe the development of the PONScore, a novel clinical tool that can be used in trials of DIPG to compare the impact of different treatments on neurological integrity. Given the successful administration by non-specialist paediatric nurses, it is our expectation that (with adequate training) the PONScore could be used outside of intensive care in a wider DIPG setting: including neurosurgery, neurology and hemato-oncology. This could be used alongside traditional methods of neurological assessment to establish pre-treatment baseline, confirm deterioration or response following treatment. Further validation in a clinical trial is required.

Acknowledgments We pay tribute to the bravery of the children and families involved in this treatment. We are grateful to Professor Steven Gill, North Bristol NHS Trust and Harley Street Clinic Children's Hospital, for his pioneering work regarding multi-catheter awake pontine infusion. Thank you to Professor David Walker, Dr Steve Lowis, Mr Peter Collins, Dr Ali Bienemann, Mr Will Singleton, Mr Neil Barua, Dr David Cronin, Dr James Morgan and Miss Marina Pitsika for their important role in developing this clinical treatment. We are thankful to Dr Kim Bull for her thoughtful insights that helped in the score's development.

Author contributions MH conducted the work, collected and analysed the data and wrote the manuscript. SZ supervised the work and developed the manuscript.

Funding Milo Hollingworth was supported with funds from the Gatsby Foundation.

Code availability Not applicable.

Compliance with ethical standards

Conflict of interest The authors report no conflict of interest.

Ethics approval Ethics approval for treatment on compassionate grounds was granted by the institutional ethics committee.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Veldhuijzen van Zanten SE, van Meerwijk CL, Jansen MH et al (2016) Palliative and end-of-life care for children with diffuse intrinsic pontine glioma: results from a London cohort study and international survey. Neuro Oncol 18:582–588
2. Hargrave D, Bartels U, Bouffet E (2016) Diffuse brainstem glioma in children: critical review of clinical trials. Lancet Oncol 7:241–248
3. Loveson KF, Fillmore HL (2018) Intersection of brain development and paediatric diffuse midline gliomas: potential role of microenvironment in tumour growth. Brain Sci 8:200
4. Cohen KJ, Jones A, Raabe E, Pearl M (2014) Highly selective intra-arterial chemotherapy for the treatment of progressive diffuse intrinsic pontine gliomas (DIPG). Neuro Oncol. https://doi.org/10.1093/neuonc/nou208.24
5. Singleton WGB, Barua NU, Morgan J et al (2016) NS-21: multi-catheter intermittent convection-enhanced delivery of carboplatin as a treatment for diffuse intrinsic pontine glioma (DIPG): pre-clinical rationale and early clinical experience. Neuro Oncol. https://doi.org/10.1093/neuonc/now078.21
6. Souweidane MM, Kramer K, Pandit-Taskar N et al (2018) Convection-enhanced delivery for diffuse intrinsic pontine glioma: a single-centre, dose-escalation, phase 1 trial. Lancet Oncol 19(8):1040–1050
7. Hollingworth M, Hurter C, Wooley M, Lewis O, Gill S, Zacharoulis S (2018) DIPG-65. preliminary experience of chronic intermittent convection enhanced delivery of carboplatin and valproic acid for the treatment of diffuse intrinsic pontine glioma following radiation therapy. Neuro Oncol. https://doi.org/10.1093/neuonc/noy059.158
8. Kieran MW (2015) Time to rethink the unthinkable: upfront biopsy of children with newly diagnosed diffuse intrinsic pontine glioma (DIPG). Pediatr Blood Cancer 62:3–4

9. Hennika T, Becher OJ (2016) Diffuse intrinsic pontine glioma: time for cautious optimism. J Child Neurol 31:1377–1385
10. Ostrom QT, Gittleman H, Farah P et al (2013) CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2006–2010. Neuro Oncol. https://doi.org/10.1093/neuonc/not151
11. Bobo RH, Laske DW, Akbasak A et al (1994) Convection-enhanced delivery of macromolecules in the brain. Proc Natl Acad Sci USA 91:2076–2080
12. Miranpuri G, Hinchman A, Wang A et al (2013) Convection enhanced delivery: a comparison of infusion characteristics in ex vivo and in vivo non-human primate brain tissue. Ann Neurosci 20:108–114
13. Killick-Cole CL, Singleton WGB, Bienemann AS et al (2017) Repurposing the anti-epileptic drug sodium valproate as an adjuvant treatment for diffuse intrinsic pontine glioma. PLoS ONE 12(5):e0176855
14. Barua NU, Lowis SP, Woolley M, O'Sullivan S, Harrison R, Gill SS (2013) Robot-guided convection-enhanced delivery of carboplatin for advanced brainstem glioma. Acta Neurochir (Wien) 155:1459–1465
15. Barua NU, Hopkins K, Woolley M et al (2016) A novel implantable catheter system with transcutaneous port for intermittent convection-enhanced delivery of carboplatin for recurrent glioblastoma. Drug Deliv 23:167–173
16. Ichord RN, Bastian R, Abraham L et al (2011) Interrater reliability of the Pediatric National Institutes of Health Stroke Scale (PedNIHSS) in a multicenter study. Stroke 42:613–617
17. Schmitz-Hubsch T, de Montcel ST, Baliko L et al (2006) Scale for the assessment and rating of ataxia: development of a new clinical scale. Neurology 66:1717–1720
18. Brandsma R, Spits AH, Kuiper MJ et al (2014) Ataxia rating scales are age-dependent in healthy children. Dev Med Child Neurol 56:556–563
19. Wong DL, Baker CM (1998) Pain in children: comparison of assessment scales. Pediatr Nurs 14:9–17
20. House JW, Brackmann DE (1985) Facial nerve grading system. Otolaryngol Head Neck Surg 93:146–147
21. Lazarus CL, Husaini H, Jacobson AS et al (2014) Development of a new lingual range-of-motion assessment scale: normative data in surgically treated oral cancer patients. Dysphagia 29:489–499
22. Campbell W, DeJong RN (2005) Cranial nerves. In: Campbell W (ed) DeJong's the neurologic examination, 7th edn. Lipincott Williams and Wilkins, Philidelphia, pp 146–161
23. Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. Stat Methods Med Res 8:135–160
24. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL (2018) Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health 6:149
25. Royal College of Nursing (2017) Standards for assessing measuring and monitoring vital signs in infants, children and young people. https://www.rcn.org.uk/-/media/royal-college-of-nursing/documents/publications/2017/may/pub-005942.pdf. Accessed 30 June 2020
26. Nayak L, DeAngelis LM, Brandes AA, Peereboom DM, Galanis E, Nancy U, Lin NU et al (2017) The Neurologic Assessment in Neuro-Oncology (NANO) scale: a tool to assess neurologic function for integration into the Response Assessment in Neuro-Oncology (RANO) criteria. Neuro Oncol 19:625–635
27. Helfer JL, Wen PY, Blakeley J, Gilbert MR, Armstrong TS (2016) Report of the Jumpstarting Brain Tumor Drug Development Coalition and FDA clinical trials clinical outcome assessment endpoints workshop (October 15, 2014, Bethesda MD). Neuro Oncol. https://doi.org/10.1093/neuonc/nov270
28. Armstrong TS, Bishof AM, Brown PD, Klein M, Taphoorn MJ, Theodore-Oklota C (2016) Determining priority signs and symptoms for use as clinical outcomes assessments in trials including patients with malignant gliomas: panel 1 report. Neuro Oncol. https://doi.org/10.1093/neuonc/nov267