Review Article

Proton versus Photon Radiotherapy for Pediatric Central Nervous System Malignancies: A Systematic Review and Meta-Analysis of Dosimetric Comparison Studies

Roberta Carbonara [ID],[1] **Alessia Di Rito,**[2] **Angela Monti,**[1] **Giuseppe Rubini** [ID],[3] **and Angela Sardaro**[1]

[1]*Interdisciplinary Department of Medicine, Section of Radiology and Radiation Oncology, University of Bari Aldo Moro, Bari, Italy*
[2]*Radiation Oncology Unit, IRCCS Istituto Tumori "Giovanni Paolo II", Bari, Italy*
[3]*Interdisciplinary Department of Medicine, Section of Nuclear Medicine, University of Bari Aldo Moro, Bari, Italy*

Correspondence should be addressed to Roberta Carbonara; roberta.carbonara@yahoo.it

Background. Radiotherapy (RT) plays a fundamental role in the treatment of pediatric central nervous system (CNS) malignancies, but its late sequelae are still a challenging question. Despite developments in modern high-conformal photon techniques and proton beam therapy (PBT) are improving the normal tissues dose-sparing while maintaining satisfactory target coverage, clinical advantages supporting the optimal treatment strategy have to be better evaluated in long-term clinical studies and assessed in further radiobiological analyses. Our analysis aimed to systematically review current knowledge on the dosimetric advantages of PBT in the considered setting, which should be the basis for future specific studies. *Materials and Methods*. A PubMed and Google Scholar search was conducted in June 2019 to select dosimetric studies comparing photon versus proton RT for pediatric patients affected by CNS tumors. Then, a systematic review and meta-analysis according to the PRISMA statement was performed. Average and standard deviation values of Conformity Index, Homogeneity Index, and mean and maximum doses to intracranial and extracranial organs at risk (OARs) were specifically evaluated for secondary dosimetric comparisons. The standardized mean differences (SMDs) for target parameters and the mean differences (MDs) for OARs were summarized in forest plots ($P < 0.05$ was considered statistically significant). Publication bias was also assessed by the funnel plot and Egger's regression test. *Results*. Among the 88 identified papers, a total of twelve studies were included in the meta-analysis. PBT showed dosimetric advantages in target homogeneity (significant especially in the subgroup comparing PBT and 3D conformal RT), as well as in the dose sparing of almost all analyzed OARs (significantly superior results for brainstem, normal brain, and hippocampal dose constraints and for extracranial OARs parameters, excluding the kidneys). Publication bias was observed for Conformity Index. *Conclusion*. Our analysis supports the evidence of dosimetric advantages of PBT over photon RT, especially in the dose sparing of normal growing tissues. Confirmations from wider well-designed studies are required.

1. Introduction

Pediatric central nervous system (CNS) malignancies are rare tumors [1] which can arise in different sites of the CNS. In recent years, patients' survival is being increased because of the advances in standard treatments [1, 2]. Radiotherapy (RT) represents a fundamental part of the recommended

multimodal therapeutic approaches, even if its late toxicity is still a question of concern in this long-surviving population [3]. In particular, cognitive and endocrine late sequelae are the most common radiation-induced side effects (RISEs) in pediatric patients treated for brain tumors [4]. Furthermore, these children are at increased risk of hearing and visual injuries, as well as vascular diseases and secondary malignant

neoplasms (SMNs), depending on the tumor site [4]. Patients treated with craniospinal irradiation (CSI) have reported a decrease in bony growth and damages to extracranial normal organs (such as the lungs and heart) [4].

Technological advances in RT planning and delivery are reducing the exposure of normal tissues, leading to improve toxicity outcomes [5]. Besides continuous advances in image-guided (IG) intensity-modulated (IM) photon radiotherapy, particle therapy with protons is establishing itself as a high-conformal RT modality which is able to improve normal tissues dose sparing while maintaining excellent target coverage [5]. Indeed, thanks to the physical characteristics of protons—such as the typical dose distribution within the "Bragg peak" [6], Proton Beam Therapy (PBT) could represent a safe alternative to photon RT for pediatric tumors or other neoplasms arising next to critical OARs [1]. Nevertheless, radiobiological uncertainties about the interaction of these charged particles with normal and neoplastic cells still persist [6].

A review of dosimetric and toxicity modeling for pediatric medulloblastoma [6] (which compared proton versus photon CSI) confirmed consistent improvements in dose sparing of out-of-field OARs and in reducing the risk of RISEs and SMNs with PBT. Nevertheless, the authors highlighted the lack of evidences from randomized prospective trials and the necessity of appropriate studies with long-term follow-up [6].

Indeed, in the past years, the limited diffusion of PBT centers and the costs of PBT treatments interfered with the account of high-level evidences from large cohorts of patients with long-term follow-up [7]. Nowadays, an increased interest in PBT is supporting its clinical application: ongoing research programs will produce higher-quality data [8].

With the aim to update the knowledge on the dosimetric advantages of PBT in the treatment of pediatric CNS tumors, we performed a systematic review and meta-analysis of published dosimetric studies that compared dosimetric outcomes between PBT and photon RT. The goal is to highlight the main emerging issues in this context that should promote specific researches with the aim to introduce advantages in clinical practice, while supporting clinical data that are being collected.

2. Materials and Methods

2.1. Search Strategy and Inclusion Criteria.
An advanced PubMed search was carried out to answer to the following research question: "What significant advantages for target and OARs dosimetry does PBT provide over photon RT in pediatric treatments for CNS tumors?"

Hence, multiple independent search strategies were performed using the following keywords (in all fields) and arrangements: (Pediatric CNS neoplasms) AND (Proton beam therapy) AND (Radiation therapy) AND (1: Brainstem dose/ 2: Cochlea dose/ 3: Optic chiasm dose/ 4: Hippocampus dose/ 5: Normal brain dose/ 6: Pituitary Gland dose/ 7: Lens dose/ 8: Retina dose/ 9: Lacrimal gland dose/ 10: Circle of Willis dose). To identify studies assessing dosimetric differences for other extracranial OARs in proton versus photon craniospinal irradiation, the keywords (Proton Craniospinal irradiation) AND (Photon Craniospinal irradiation) AND (Dosimetry OR Dosimetric study) were additionally searched.

Searches were completed in June 2019. To identify more references, no restrictions on years or publication type were considered. Indeed, to collect additional eligible studies, we searched supplementary references cited by more recent retrieved review articles. Furthermore, an additional search in Google Scholar was performed for analogous purposes.

A systematic review according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [9] was independently conducted by two authors. Study selection criteria—including screening and eligibility requirements—are reported in Table 1. All studies satisfying the eligibility criteria were included in qualitative and quantitative synthesis.

2.2. Data Extraction.
We collected and analyzed all useful dosimetric data which were provided by the eligible papers for both target volumes and OARs, regardless of the specific search strategies adopted for study selection.

The following basic data were extracted from the included studies: first author name, publication year, tumor histology, sample size, study assessment, and total target dose.

Mean (Dmean) and maximum (Dmax) doses expressed in Gy were specifically considered for our secondary analyses. Whenever possible, it was expected to convert the reported relative values (%) of mean and maximum doses into the corresponding absolute values (Gy). For comparative purposes, average and standard deviation values of Dmean and Dmax were extracted by papers or calculated if raw data were available. If all these data were not available, then the paper was not included in the qualitative synthesis and meta-analysis.

For photon treatments, if the articles provided data by both linac and tomotherapy, we reviewed data of linac-based treatments because of their larger utilization. When both IMRT and VMAT plans were assessed, we extracted and analyzed VMAT data. Similarly, for proton treatments, if the articles provided data by both passively scattered/3D conformal proton therapy and scanning/intensity-modulated (IM) proton therapy, we specifically evaluated data from these latter techniques because of their superior plan quality, as reported in previous published works [1].

2.3. Statistical Analysis.
We calculated the standardized mean differences (SMDs) with a 95% confidence interval (CI) between photon and proton plans for target dosimetric parameters (Homogeneity Index and Conformity Index). The mean differences (MDs) of Dmax and Dmean values were also calculated between the considered RT modalities with the respective 95% CI.

I^2 was used to assess heterogeneity between studies. If heterogeneity was not present ($I^2 < 50\%$), a fixed-effect model was performed for our analysis; otherwise, a random-effect model was adopted. $P < 0.05$ was considered statistically significant.

Table 1: Study selection criteria.

Inclusion criteria	Exclusion criteria
Screening criteria: population Pediatric and young adult patients (age < 21 years) affected by CNS neoplasms (craniopharyngioma, ependymoma, neuroblastoma, CNS germinoma, glioma, medulloblastoma, and primitive neuroectodermal tumors (PNETs))	Mixed populations (adults and pediatrics), or adults (age > 21) Non-CNS malignancies
Screening criteria: study design Dosimetric comparison between proton beam therapy and photon radiotherapy Dosimetric studies comparing the most advanced/widespread irradiation techniques (protons versus photons)	Studies reporting single techniques and other particle therapy modalities Reviews, clinical case reports/case series, cost-effective studies, simulation studies, preclinical models, etc. Letters, editorials, congress abstracts, and guidelines
Screening criteria: outcomes Dosimetric results for target, intracranial OARs (brainstem, cochlea, optic chiasm, hippocampus, normal brain tissue, pituitary gland, and Circle of Willis), and extracranial OARs (lens, retina, lacrimal glands, thyroid, esophagus, lungs, heart, liver, and kidneys)	Absence of reporting of dosimetric outcomes related to the target and/or the considered OARs
Eligibility criteria: outcome measures Studies reporting any of the following parameters with average and standard deviation values: (i) For OARs: Dmax for the brainstem, optic chiasm, normal brain tissue, pituitary gland, lens, retina, lacrimal gland, and esophagus and Dmean for the brainstem, cochlea, optic chiasm, hippocampus, normal brain tissue, pituitary gland, circle of Willis, lens, lacrimal gland, and other extracranial OARs (ii) For target: Homogeneity Index and Conformity Index	Studies reporting other dosimetric parameters for target and OARs
Screening criteria: language English	All other languages

Whenever possible, subgroup analyses were performed to assess differences between photon techniques (3D-CRT versus intensity-modulated techniques).

Publication bias was evaluated by visual inspection of the funnel plot and Egger's regression test. Egger's P value <0.1 was considered as significant asymmetry. All statistical analyses were performed using Review Manager (RevMan) version 5.3 and Comprehensive Meta-Analysis version 3.0.

3. Results

A total of 88 papers were identified from different sources. PubMed results according to specific keywords were as follows: *brainstem dose n* = 11, *cochlea dose n* = 9, *optic chiasm dose n* = 5, *hippocampus dose n* = 4, *normal brain dose n* = 27, *pituitary gland dose n* = 5, *lens dose n* = 2, and *circle of Willis dose n* = 1; for CSI dosimetric studies, total results *n* = 21. Additional searches on the retrieved articles and Google Scholar provided 4 results. Figure 1 shows the study flow chart according to the PRISMA statement [9]. Finally, 12 studies were eligible for inclusion in our meta-analysis (Figure 1).

Basic characteristics of included studies are summarized in Table 2. Forest plots for target parameters and OARs doses are shown in Figures 2–5.

Data from single studies—which cannot be aggregated in forest plots for a quantitative synthesis—are summarized in a table submitted as a supplementary material. In case of too small sample size and heterogeneity in the definition of anatomical structures, the results were synthetized in a qualitative manner.

Six studies [10–15]—which included a total of 72 patients—were evaluated for Homogeneity Index assessment. A significant overall advantage in homogeneity of target dose distribution was observed with PBT (SMD: 0.90, 95% CI: 0.02, 1.78, P = 0.04), with a major improvement in the 3D-CRT subgroup (SMD: 3.40, 95% CI: 1.93, 4.87, P < 0.00001) (Figure 2). Nevertheless, no significant differences in the IMRT/VMAT subgroup were observed (P = 0.16) (Figure 2).

Among the three analyzed studies which provided data for Conformity Index [10, 12, 16], no significant differences were found between the RT modalities (P = 0.14), even if significant superior results were reported with protons by Beltran et al. [16] and Freund et al. [12] (Figure 2).

Meta-analyses of intracranial OARs mean doses (Figure 3) showed significantly improved results with protons for the following organs: brainstem (MD: 2.07, 95% CI: 1.21, 2.93, P < 0.00001), right hippocampus (MD: 5.71, 95% CI:

Total identified studies
(PubMed source)
$n = 85$

Records after duplicates removed (18)
$n = 67$

Records screened
(title/abstract evaluation)
$n = 67$

Records excluded ($n = 45$):
articles discussing non-CNS tumors and clinical,
technological, or radiobiological issues;
single-arm case series; and model studies, review
articles, abstract, and consensus guidelines

Full-text articles assessed
for eligibility
$n = 22$

Full-text articles excluded, $n = 14$,
with reasons: studies not reporting average
and SD values; single-patient plan
evaluation; and other dose level parameters
evaluated

Studies included in
qualitative synthesis
$n = 8$

Studies identified through other sources
(retrieved articles/Google Scholar)
$n = 4$

Studies included in quantitative
synthesis (meta-analysis)
$n = 12$

Figure 1: Flow chart of the research strategy according to the PRISMA statement [9].

0.25, 11.16, $P = 0.04$), normal brain (MD: 5.08, 95% CI: 3.36, 6.80, $P < 0.00001$), and optic chiasm (MD: 4.32, 95% CI: 2.37, 6.28, $P < 0.00001$). PBT also showed improved results in the studies reporting mean doses to the left and right cochlea, left hippocampus, and pituitary gland, even if these improvements did not provide overall significant differences as compared to photon RT (Figure 3).

Meta-analysis of brainstem maximum doses (Figure 4) revealed a significant advantage with protons ($P = 0.02$), while no significant difference emerged for the normal brain maximum dose ($P = 0.63$) between the analyzed RT modalities (Figure 4).

Globally, three studies [11, 14, 17] provided useful dosimetric data for extracranial OARs of the cephalic district: mean left lens dose, maximum left lens doses, and mean doses to lacrimal glands were significantly improved in PBT plans (Figure 5). Three supplementary studies [15, 18, 19] assessed dosimetry of cervical, thoracic, and abdominal OARs: mean doses to the esophagus, thyroid, lungs, and liver were significantly improved with PBT, while no significant overall advantage for the kidneys was observed ($P = 0.11$) (Figure 5). Nevertheless, subgroup analyses according to photon RT techniques performed for the kidneys showed significant superior results with protons over both intensity-

modulated and 3D conformal photon techniques ($P < 0.00001$ in both cases) (Figure 5). The mean difference in the IMRT subgroup was higher than that in the 3D-CRT group: 7.60 (95% CI: 6.98, 8.22) versus 1.47 (95% CI: 1.04, 1.89). A similar higher result in the IMRT subgroup was found for the lungs, as opposed to the results of subgroup analyses performed for the esophagus (Figure 5).

Single studies reporting dosimetric comparisons for the heart mean dose [19], right lens mean dose [14], and optic chiasm and pituitary and lacrimal gland maximum doses [11] substantially confirmed the advantages of PBT over photon RT (see Supplementary Materials) (available here), with the exception of pituitary gland Dmax in Correia's study [11]. Boehling compared mean and maximum doses to vascular structures of the circle of Willis between PBT (with both 3D-PBT and IMPT) and IMRT, confirming the dosimetric advantages of protons [10]. In the study by Correia, only the mean dose to the circle of Willis was reduced by PBT as compared to IMRT and VMAT [11].

3.1. Publication Bias. Visual evaluation of funnel plots and Egger's tests were performed for meta-analyses that included at least three studies. Thus, we were able to evaluate

TABLE 2: Dosimetric studies assessing Conformity Index, Homogeneity Index, and Dmean and/or Dmax for OARs.

Authors (year)	Tumor histology	Patient number	Dosimetric study assessment	Mean total target dose (Gy/RBE/CGE) (dose/fraction)	Evaluation of at least one target parameter: CI (or CN), HI	Dmean/Dmax for OARs (Gy or %) with mean and SD	Conclusions
Stoker et al. (2018) [14]	Primary brain tumors requiring hippocampal-avoidance- (HA-) WBRT	10/20	Dosimetric comparison between VMAT and IMPT for HA-WBRT	36 Gy (1.8 Gy/die) HA-WBRT	HI	Dmax and Dmean reported for the normal brain, hippocampi, cochlea, and lens and Dmean for the brainstem	HA-IMPT can match or improve dosimetric benefits obtained with VMAT.
Freund et al. (2015) [12]	Glioma Ependymoma	8 5	Dosimetric comparison between VMAT, PSPT, and IMPT and risk of cerebral radionecrosis assessment	54 Gy (RBE) (1.8 Gy/die)	CI, HI	Dmax and Dmean evaluated and reported for the normal brain	Both PSPT and IMPT plans significantly improved the maximum dose to the brain. A significant lower risk of brain radionecrosis was observed with PBT.
Howell et al. (2012) [13]	Medulloblastoma	18	Comparison of dose distributions and DVHs between photon and proton CSI	23.4 Gy (1.8 Gy/fr)	CI, HI	Dmean and/or Dmax not reported for the analyzed OARs	Both photon and proton plans provided good target coverage; PBT dose distributions were more homogeneous. Proton CSI improved normal tissue sparing.
Correia et al. (2019) [11]	Intracranial germ-cell tumor	11	Comparison of dose distributions and DVHs between WV-RT/TB IMRT, VMAT, and PBS-PT	24 Gy (RBE) WV-RT plus boost up to 40 Gy (1.6 Gy/fr)	HI and inhomogeneity coefficient	Dmean and Dmax reported (%) for the brainstem, chiasm, normal brain, pituitary gland, circle of Willis, bilateral cochlea, hippocampus, lens, and lacrimal gland	PBS-PT was superior to photons in conformality and OAR sparing.
Boehling et al. (2012) [10]	Craniopharyngioma	10	Comparison of dose distributions and DVHs between IMRT, 3D-PRT, and IMPT	50.4 Gy (CGE) (1.8 Gy/fr)	CN, HI	Dmean and Dmax reported for the vascular OARs, brainstem, and normal brain	PBT was able to avoid excess integral dose to a variety of normal structures at all dose levels while maintaining equal target coverage.

TABLE 2: Continued.

Authors (year)	Tumor histology	Patient number	Dosimetric study assessment	Mean total target dose (Gy/ RBE/CGE) (dose/fraction)	Evaluation of at least one target parameter: CI (or CN), HI	Dmean/Dmax for OARs (Gy or %) with mean and SD	Conclusions
Takizawa et al. (2017) [31]	Ependymoma Germinoma	6 6	Comparison of dose distributions and DVHs between PBT, 3D-CRT, and IMRT	Median of 52.2 Gy for ependymoma and median of 30.6 Gy for germinoma	Not reported	Normal brain dose reported for each patient and as a percentage of the prescription dose (visual inspection of raw data)	PBT reduces the average dose to normal brain tissue as compared to 3D-CRT and IMRT.
MacDonald et al. (2008) [26]	Ependymoma	2/17	Comparison of dose distributions and DVHs between IMPT, 3D-PBT, and IMRT	55.8 Gy	Not reported	Dmean for the brain, brainstem, pituitary gland, optic chiasm, and cochlea evaluated and reported for each patient (Gy)	Dose distributions for PBT were compared favourably with IMRT plans. IMPT allows further sparing of some critical structures.
Beltran et al. (2012) [16]	Craniopharyngioma	14	Dosimetric comparison between IMRT, double-scatter (DS) PT, and IMPT	54 Gy (1.8 Gy/ die)	CI	Not reported (other dosimetric parameters are reported)	PBT significantly reduced the dose to the whole brain. IMPT was the most conformal treatment that improved OAR dose sparing, but it was highly sensitive to target changes.
Dennis et al. (2013) [17]	Low-grade glioma	11	Dosimetric (DVH) comparison between IMRT and PBT. SMN risk assessment	54 Gy (1.8 Gy/ die)	Not reported	Dmean for the brainstem, pituitary gland, optic chiasm, and lacrimal gland evaluated and reported for each patient (Gy)	PBT improved the reduction of doses to normal tissues, especially when tumors were in close proximity to critical structures. IMRT had a twofold higher risk of SMNs as compared to PBT.
Mu et al. (2005) [18]	Medulloblastoma	5	Dosimetric comparison between conventional photons, IMRT, electrons, and PBT. SMN risk assessment	23.4 Gy (1.8 Gy/ die)	Not reported	Dmean evaluated and reported for the thyroid, esophagus, heart, lungs, and liver	IMPT significantly reduced mean doses to OARs, except for the lungs (not significantly). IMPT reduced SMN risk.

TABLE 2: Continued.

Authors (year)	Tumor histology	Patient number	Dosimetric study assessment	Mean total target dose (Gy/RBE/CGE) (dose/fraction)	Evaluation of at least one target parameter: CI (or CN), HI	Dmean/Dmax for OARs (Gy or %) with mean and SD	Conclusions
Zhang et al. (2014) [19]	Medulloblastoma	17	Dosimetric comparison between PSPT CSI and field-in-field photon CSI. SMN risk assessment	23.4 Gy (1.8 Gy/die)	Not reported	Dmean evaluated and reported for the thyroid, heart, lungs, and liver	PSPT CSI provided lower doses to OARs, superior predicted outcomes, and lower predicted risks of SMNs and cardiac mortality than field-in-field photon CSI.
Yoon et al. (2011) [15]	Various CNS tumors	10	Comparison of dose distributions, DVHs, and SMN risk between CSI with 3D-CRT, TOMO, and PBT. SMN risk assessment	36 Gy (1.8 Gy/fr) to the spine; total target dose ranged between 54 and 60.6 Gy	CI, HI	Dmean evaluated and reported for the lens, thyroid, esophagus, lungs, liver, and kidneys	PBT provided the best HI and a superior CI than 3D-CRT (no significant difference compared to TOMO). OAR doses with PBT were lower than those obtained with 3D-CRT or TOMO. Lower SMN risk was reported with PBT.

CI: Conformity Index; HI: Homogeneity Index; CGE: cobalt Gy equivalents; RBE: relative biological effectiveness; SD: standard deviation; CSI: craniospinal irradiation; TOMO: tomotherapy; PBS-PT: pencil beam scanning-proton therapy; PSPT: passively scattered PT; VMAT: volumetric modulated arc therapy; IMRT/IMPT: intensity-modulated radiotherapy or PT; SMNs: secondary malignant neoplasms; WBRT: whole-brain RT; WV-RT/TB: whole-ventricular RT followed by a boost to the tumor bed.

publication bias for the following parameters (Figures 6 and 7): target homogeneity and conformity, Dmean of the brainstem, optic chiasm, left and right cochlea, normal brain, esophagus, lungs, and kidneys, and Dmax of the normal brain.

The funnel plot of Homogeneity Index appeared symmetrical, and these findings were confirmed by Egger's regression tests ($P = 0.11$), while a significant asymmetry was found for Conformity Index ($P = 0.0088$). No significant publication bias was found for all the analyzed OARs (Figures 6 and 7): brainstem ($P = 0.24$), optic chiasm ($P = 0.74$), left and right cochlea ($P = 0.28$ and $P = 0.46$, respectively), normal brain (Dmean $P = 0.4$, Dmax $P = 0.89$), esophagus ($P = 0.99$), lungs ($P = 0.61$), and kidneys ($P = 0.85$).

4. Discussion

4.1. Radiotherapy in Pediatric CNS Tumors. The 2016 World Health Organization (WHO) classification of CNS tumors [20] emphasizes a huge variety of these neoplasms due to their phenotypical and molecular characteristics which reflect the genetic basis of tumorigenesis.

A variety of tumor histology, grades, and primary locations require different RT prescriptions to provide a radical or adjuvant local disease management. Indeed, different RT treatment fields and doses are used in clinical practice, varying from CSI plus a boost for medulloblastoma/primitive neuroectodermal tumors (PNETs) (depending on the risk level, 24 or 36 Gy could be delivered to the craniospinal axis, followed by a boost to the posterior fossa or to the tumor bed (up to a total dose not inferior to 54 Gy) [4, 21]) and some cases of germ-cell tumors [4, 22] to a resected tumor bed irradiation—e.g., for high-grade glioma, ependymoma (that receives 54 Gy followed by a boost up to 59.4 Gy [4, 22]), craniopharyngioma (prescription doses between 45 and 59.4 Gy have been reported [4, 22]), and some cases of germ-cell tumors [4]—or a whole-ventricular (WV) RT with or without a localized boost (up to 50–54 Gy) for some cases of germ-cell tumors [4, 22].

Previous published authoritative literature [22] has summarized the dosimetric advantages of protons over photons in radiation treatments for pediatric CNS tumors. To our knowledge, to date, this is the first literature review that provides a meta-analysis of dosimetric comparison studies to systematize PBT dosimetric outcomes.

Study or subgroup	Photon			Proton			Weight	Std. mean difference IV, random, 95% CI
	Mean	SD	Total	Mean	SD	Total		
Boehling, 2012	0.76	0.07	10	0.73	0.05	10	34.0%	0.47 [−0.42, 1.36]
Beltran, 2012	0.71	0.04	14	0.84	0.04	14	33.0%	−3.16 [−4.31, −2.00]
Freund, 2015	0.62	0.08	13	0.83	0.06	13	33.0%	−2.88 [−4.02, −1.73]
Total (95% CI)			37			37	100.0%	−1.83 [−4.28, 0.62]

Heterogeneity: tau^2 = 4.40; chi^2 = 32.12; df = 2 (P < 0.00001); I^2 = 94%
Test for overall effect: Z = 1.46 (P = 0.14)

(a)

Study or subgroup	Photon			Proton			Weight	Std. mean difference IV, random, 95% CI
	Mean	SD	Total	Mean	SD	Total		
1.1.1 IMRT/VMAT								
Boehling, 2012	0.25	0.06	10	0.23	0.06	10	17.2%	0.32 [−0.56, 1.20]
Correia, 2019	5	1.4	11	4.7	1	11	17.5%	0.24 [−0.60, 1.08]
Freund, 2015	0.07	0.05	13	0.12	0.12	13	17.8%	−0.53 [−1.31, 0.26]
Howell, 2012	1.05	0.009	18	1.04	0.012	18	18.4%	0.92 [0.23, 1.61]
Stoker, 2018	15.1	6.3	10	6.8	0.8	10	15.9%	1.77 [0.70, 2.84]
Subtotal (95% CI)			62			62	86.8%	0.51 [−0.19, 1.21]

Heterogeneity: tau^2 = 0.45; chi^2 = 13.93; df = 4 (P = 0.008); I^2 = 71%
Test for overall effect: Z = 1.42 (P = 0.16)

1.1.2 3D-CRT								
Yoon, 2011	28.7	5.9	10	12.8	2.3	10	13.2%	3.40 [1.93, 4.87]
Subtotal (95% CI)			10			10	13.2%	3.40 [1.93, 4.87]

Heterogeneity: not applicable
Test for overall effect: Z = 4.54 (P < 0.00001)

Total (95% CI)			72			72	100.0%	0.90 [0.02, 1.78]

Heterogeneity: tau^2 = 0.98; chi^2 = 28.43; df = 5 (P < 0.0001); I^2 = 82%
Test for overall effect: Z = 2.01 (P = 0.04)
Test for subgroup differences: chi^2 = 12.16; df = 1 (P = 0.0005); I^2 = 91.8%

(b)

FIGURE 2: Target conformity and homogeneity. (a) Conformity Index. (b) Homogeneity Index.

The choice to specifically evaluate data from high-conformal PBT techniques (such as IMPT) when they were available—even if it could have limited the analysis to not very widespread PBT modality—was aimed to improve the knowledge on outcomes of advanced technologies in PBT treatments. The expected findings could support more sophisticated PBT planning and become relevant for medical physicists, medical dosimetrists, and radiation oncologists in the next years.

Similarly, the choice to compare these findings with data from available high-conformal photon techniques was aimed at a preliminary comparative analysis which should encourage further studies, also including cost-effective comparative analyses.

4.2. Target Dosimetry Assessment.

We compared target dosimetry between proton and photon plans based on the Homogeneity Index and Conformity Index. The first parameter is used to quantify the homogeneity of dose distribution within the target volume, while the second one is used to quantify the conformation of the prescribed dose to the target volume. Superior results in target dosimetry are established according to the highest target conformity (highest value of Conformity Index) and homogeneity (lowest value of Homogeneity Index) [10, 13, 23].

Target conformity and homogeneity were acceptable in all analyzed dosimetric studies. Globally, while homogeneity was significantly improved with PBT—with higher

significance in the study by Yoon et al. on CSI [15], confirming the expected advantage of PBT over 3D-CRT [12]—our results showed no significant differences in target conformity and homogeneity between protons and high-conformal photon techniques (IMRT/VMAT) (Figure 2).

The use of high-conformal photon RT has been introduced in clinical practice with the primary aim to improve the dose sparing of normal tissues [23], but these modern techniques substantially improve target conformity and homogeneity, even if major prerequisites for IMRT and VMAT remain the adequate delineation of target volumes and the management of target motion [23]. It is reasonable that the advantages in dose distribution could lead to a safety delivery of higher doses to target volumes, thus improving treatment efficacy.

While we were reaching for useful clinical correlations among the analyzed studies, we observed that MacDonald reported that PBT performed for ependymoma patients was compared favourably with the literature for disease control outcomes [26] over a median follow-up of 26 months.

Confirmations from adequate clinical and radiobiological studies (that should take into account the interaction between protons and tumor cells) are required to clarify the advantage of PBT in tumor control. Results from appropriate cost-effective analyses comparing high-conformal photon RT and PBT will also support the correct management of pediatric CNS tumors.

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, fixed, 95% CI
Correia, 2019	30.92	2.47	11	29.92	2.15	11	19.6%	1.00 [−0.94, 2.94]
Boehling, 2012	25.2	10.7	10	21	10	10	0.9%	4.20 [−4.88, 13.28]
MacDonald, 2008 (1)	31	8	2	22.5	1.5	2	0.6%	8.50 [−2.78, 19.78]
Dennis, 2013	36.36	11.86	11	24.27	16.04	11	0.5%	12.09 [0.30, 23.88]
Stoker, 2018	38	1.2	10	35.8	1	10	78.4%	2.20 [1.23, 3.17]
Total (95% CI)			44			44	100.0%	2.07 [1.21, 2.93]

Heterogeneity: chi² = 5.48; df = 4 (P = 0.24); I² = 27%
Test for overall effect: Z = 4.74 (P < 0.00001)

(a)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI
Correia, 2019	5.28	0.21	11	4.8	0.38	11	50.1%	0.48 [0.22, 0.74]
MacDonald, 2008 (2)	20	17	2	1.045	0.955	2	5.8%	18.95 [−4.64, 42.55]
Stoker, 2018	37.2	1.9	10	30.6	4.7	10	44.1%	6.60 [3.46, 9.74]
Total (95% CI)			23			23	100.0%	4.25 [−1.78, 10.27]

Heterogeneity: tau² = 18.84; chi² = 16.82; df = 2 (P = 0.0002); I² = 88%
Test for overall effect: Z = 1.38 (P = 0.17)

(b)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI
Correia, 2019	5.4	0.243	11	5.32	0.58	11	52.0%	0.08 [−0.29, 0.45]
MacDonald, 2008 (1)	22.5	20.5	2	14.545	14.455	2	2.8%	7.96 [−26.81, 42.72]
Stoker, 2018	37.5	2.1	10	31	4.8	10	45.2%	6.50 [3.25, 9.75]
Total (95% CI)			23			23	100.0%	3.20 [−2.79, 9.19]

Heterogeneity: tau² = 17.94; chi² = 15.01; df = 2 (P = 0.0005); I² = 87%
Test for overall effect: Z = 1.05 (P = 0.29)

(c)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI
Correia, 2019	35.68	2.82	11	33.8	3.55	11	47.8%	1.88 [−0.80, 4.56]
Stoker, 2018	13.7	0.8	10	5.4	0.2	10	52.2%	8.30 [7.79, 8.81]
Total (95% CI)			21			21	100.0%	5.23 [−1.06, 11.52]

Heterogeneity: tau² = 19.64; chi² = 21.28; df = 1 (P < 0.00001); I² = 95%
Test for overall effect: Z = 1.63 (P = 0.10)

(d)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI
Correia, 2019	34.56	3.08	11	31.84	3.95	11	46.5%	2.72 [−0.24, 5.68]
Stoker, 2018	13.7	0.8	10	5.4	0.3	10	53.5%	8.30 [7.77, 8.83]
Total (95% CI)			21			21	100.0%	5.71 [0.25, 11.16]

Heterogeneity: tau² = 14.39; chi² = 13.23; df = 1 (P = 0.0003); I² = 92%
Test for overall effect: Z = 2.05 (P = 0.04)

(e)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, fixed, 95% CI
Correia, 2019	31.36	6.87	11	29.2	7.97	11	9.9%	2.16 [−4.06, 8.38]
Dennis, 2013	35.95	18.13	11	22.14	20.13	11	1.5%	13.81 [−2.20, 29.82]
MacDonald, 2008 (2)	4.5	1.5	2	0.095	0.005	2	88.6%	4.41 [2.33, 6.48]
Total (95% CI)			24			24	100.0%	4.32 [2.37, 6.28]

Heterogeneity: chi² = 1.82; df = 2 (P = 0.40); I² = 0%
Test for overall effect: Z = 4.33 (P < 0.0001)

(f)

FIGURE 3: Continued.

Study or subgroup	Photon Mean	Photon SD	Photon Total	Proton Mean	Proton SD	Proton Total	Weight	Mean difference IV, random, 95% CI
Correia, 2019	17.84	1.62	11	12	1.03	11	23.1%	5.84 [4.71, 6.97]
Boehling, 2012	7.6	2.7	10	5.6	1.8	10	19.0%	2.00 [−0.01, 4.01]
Takizawa, 2017 (1)	14.79	8.87	6	8.6	6.17	6	3.4%	6.19 [−2.46, 14.84]
Takizawa, 2017 (2)	18.33	2.26	6	14.88	1.89	6	17.3%	3.45 [1.09, 5.81]
MacDonald, 2008 (1)	12.5	0.5	2	5.5	0.5	2	23.7%	7.00 [6.02, 7.98]
Freund, 2015	17.5	5.6	13	10.3	4	13	11.6%	7.20 [3.46, 10.94]
Stoker, 2018	38.9	19.9	10	36.3	0.5	10	1.8%	2.60 [−9.74, 14.94]
Total (95% CI)			58			58	100.0%	5.08 [3.36, 6.80]

Heterogeneity: tau^2 = 3.00; chi^2 = 24.10; df = 6 (P = 0.0005); I^2 = 75%
Test for overall effect: Z = 5.79 (P < 0.00001)

Favours (photon)   Favours (proton)

(g)

Study or subgroup	Photon Mean	Photon SD	Photon Total	Proton Mean	Proton SD	Proton Total	Weight	Mean difference IV, random, 95% CI
Correia, 2019	21.56	8.495	11	21.8	8.982	11	52.7%	−0.24 [−7.55, 7.07]
Dennis, 2013	33.618	17.568	11	11.209	12.287	11	47.3%	22.41 [9.74, 35.08]
Total (95% CI)			22			22	100.0%	10.47 [−11.69, 32.63]

Heterogeneity: tau^2 = 228.65; chi^2 = 9.21; df = 1 (P = 0.002); I^2 = 89%
Test for overall effect: Z = 0.93 (P = 0.35)

Favours (photon)   Favours (proton)
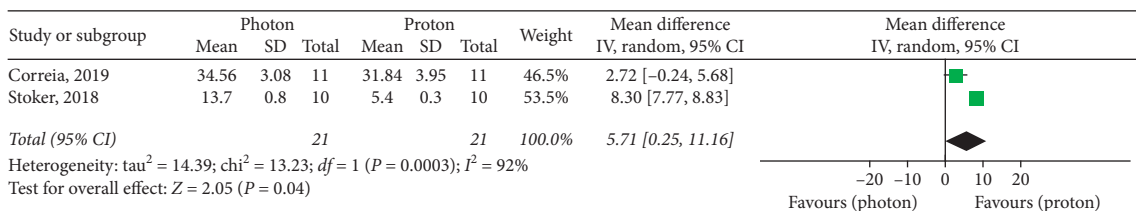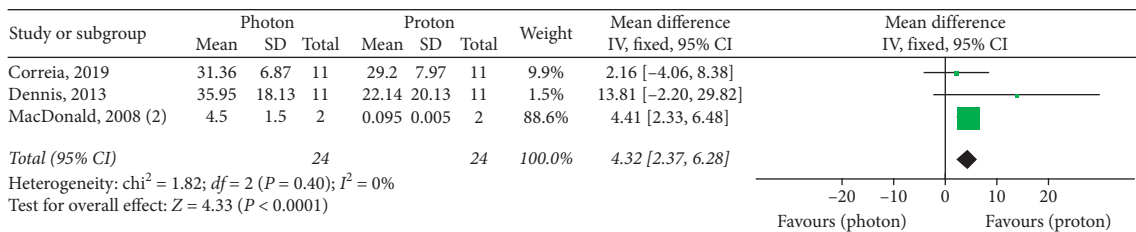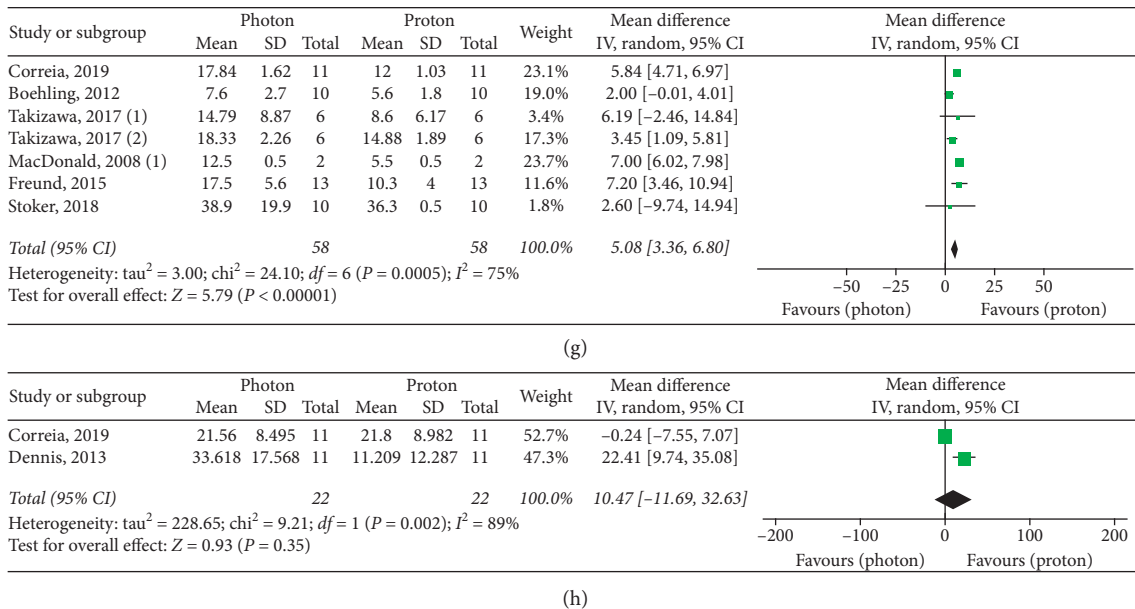
(h)

FIGURE 3: Dmean of intracranial OARs. (a) Brainstem. (b) Left cochlea. (c) Right cochlea. (d) Left hippocampus. (e) Right hippocampus. (f) Optic chiasm. (g) Normal brain. (h) Pituitary gland. When not otherwise specified, the photon RT technique is IMRT or VMAT. Takizawa, 2017 (1): patients affected by ependymoma (IMRT versus PBT); Takizawa, 2017 (2): patients affected by germinoma (IMRT versus PBT); MacDonald, 2008 (1): proton technique: IMPT; MacDonald, 2008 (2): proton technique: 3D conformal PBT.

4.3. Dosimetric Analyses for OARs.

In our meta-analyses for intracranial and extracranial OARs, we specifically evaluated dosimetric data of OARs involved in neurogenesis (hippocampus [10, 11, 27]), sensory functions (optic chiasm, lens, retina, lacrimal gland, and cochlea), endocrine functions (pituitary gland and thyroid), neurocognition (hippocampus, normal brain, and brainstem), and tissues exposed to SMNs risk (normal brain tissue and brainstem [10, 11]) or late toxicity (circle of Willis for cerebrovascular disease [10, 11] and esophagus, thyroid, heart, lungs, liver, and kidneys).

An adequate respect of intracranial OARs dose constraints (brainstem: Dmax < 54 Gy [27–29]; normal brain: Dmean < 25 Gy [28, 29]; cochlea: Dmean < 35 Gy [27] or <37 Gy [28, 29]; optic chiasm: Dmax < 55 Gy [27] or <52 Gy [28, 29]; hippocampus: Dmean < 30 Gy [27]; and pituitary gland: Dmax < 42 Gy and Dmean < 25 or 30 Gy [27]) was observed in almost all proton and photon plans, but PBT improved the dose sparing for all the analyzed structures (Figures 3 and 4), showing significantly superior results for dose constraints of the brainstem, normal brain, and hippocampus.

Furthermore, even if both proton and photon plans ensured a satisfactory dose sparing for the lacrimal gland (Dmax < 40 Gy [27] and Dmean < 20 Gy [28, 29]), lungs (Dmean < 10 Gy [28, 29]), liver (Dmean < 10 Gy [28, 29]), and kidneys (Dmean < 16 Gy [28, 29]), only PBT achieved a useful dose sparing for the lens (Dmax < 10 Gy [27]), thyroid (Dmean < 6 Gy [28, 29]), and heart (Dmean < 3.5 Gy [28, 29]) among the considered studies. Our analysis also emphasized that PBT provided an improvement in dose sparing of these extracranial OARs, which was significant in all cases except for the kidneys (Figure 5).
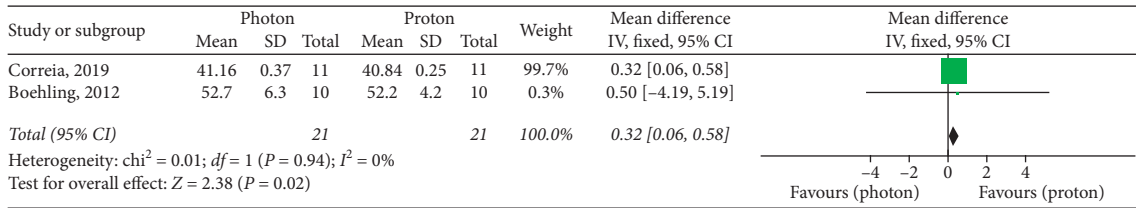
In the studies by Boehling et al. and Correia et al. [10, 11], the inappropriate doses to vascular structures of the circle of Willis (which were distinguished by Boehling in anterior and middle cerebral arteries and anterior communicating arteries [10]) were reduced in proton plans. Higher doses to these structures have been shown to correlate with vascular damages, such as Moyamoya syndrome and cerebrovascular disease (ischemic events, in primarily) [10]. Thus, in the case of dose sparing of either vascular structures [10] or other OARs [6], the dosimetric advantages achieved with protons are expected to translate into clinical benefits.

Nevertheless, besides the reduction of OARs doses, Ho et al. [6] underlined the added importance of homogeneous doses to OARs for clinical improvements.

Also, it has to be noted that, since critical structures could be enclosed within the target volume (in-field OARs) or could be in their close proximity when WBRT or WV-RT [11] or CSI [13] is performed, the normal tissues' dosimetric parameters are influenced by the OAR's location [13], as well as treatment fields and total target dose. Accordingly, the degree and extent of neurocognitive deterioration have shown to be affected by the total radiation dose [30] and tumor volume and site, as well as by the age of patient at treatment time [31, 32].

Because of the relevance of these issues in the understanding and disclosing of dosimetric differences between PBT and photon RT for OARs, we underline the heterogeneity of RT treatment fields and target doses among the included studies, which could have influenced our secondary comparisons of plan performance.

In particular, besides the studies on CSI [15, 18, 19], Takizawa et al. [31] and Correia et al. (who also planned a boost to the tumor bed) [11] analyzed whole-ventricular

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, fixed, 95% CI	Mean difference IV, fixed, 95% CI
Correia, 2019	41.16	0.37	11	40.84	0.25	11	99.7%	0.32 [0.06, 0.58]	
Boehling, 2012	52.7	6.3	10	52.2	4.2	10	0.3%	0.50 [−4.19, 5.19]	
Total (95% CI)			21			21	100.0%	0.32 [0.06, 0.58]	

Heterogeneity: chi^2 = 0.01; df = 1 (P = 0.94); I^2 = 0%
Test for overall effect: Z = 2.38 (P = 0.02)

Favours (photon) Favours (proton)

(a)

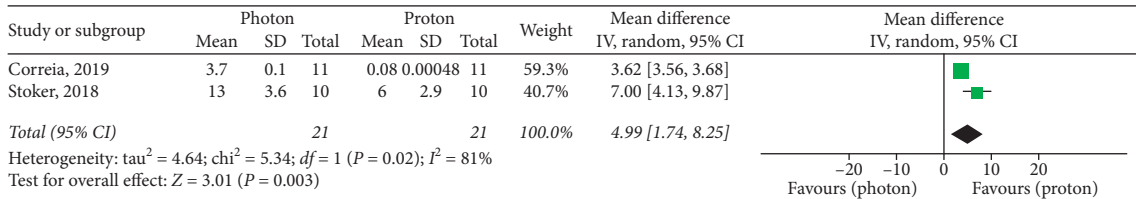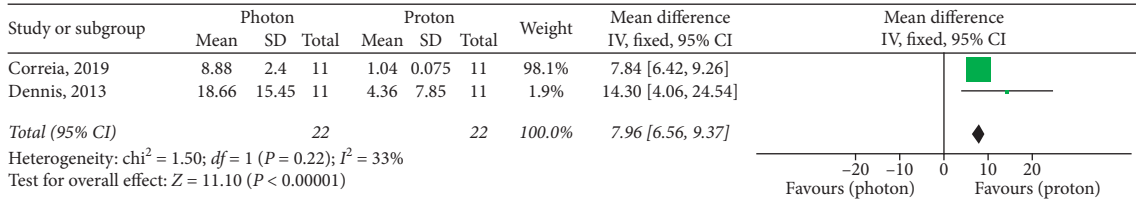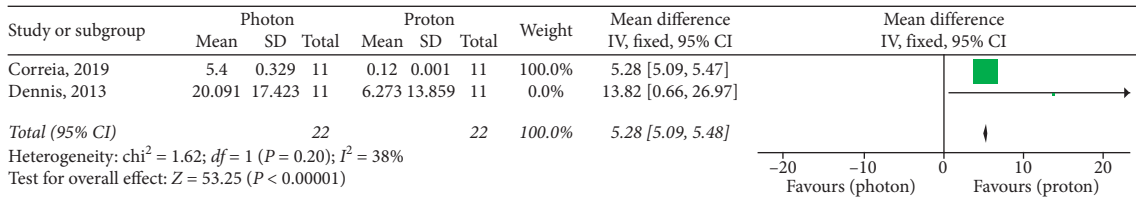| Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI	Mean difference IV, random, 95% CI
Correia, 2019	40.92	0.61	11	40.28	0.2	11	29.4%	0.64 [0.26, 1.02]	
Boehling, 2012	54.7	1.1	10	54.7	1.6	10	25.4%	0.00 [−1.20, 1.20]	
Freund, 2015	55.8	1.08	13	57.6	1.44	13	26.8%	−1.80 [−2.78, −0.82]	
Stoker, 2018	44.1	3.6	10	40.4	0.7	10	18.3%	3.70 [1.43, 5.97]	
Total (95% CI)			44			44	100.0%	0.38 [−1.18, 1.94]	

Heterogeneity: tau^2 = 2.11; chi^2 = 29.45; df = 3 (P < 0.00001); I^2 = 90%
Test for overall effect: Z = 0.48 (P = 0.63)

Favours (photon) Favours (proton)

(b)

Figure 4: Dmax of intracranial OARs. (a) Brainstem. (b) Normal brain.

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI	Mean difference IV, random, 95% CI
Correia, 2019	3.7	0.1	11	0.08	0.00048	11	59.3%	3.62 [3.56, 3.68]	
Stoker, 2018	13	3.6	10	6	2.9	10	40.7%	7.00 [4.13, 9.87]	
Total (95% CI)			21			21	100.0%	4.99 [1.74, 8.25]	

Heterogeneity: tau^2 = 4.64; chi^2 = 5.34; df = 1 (P = 0.02); I^2 = 81%
Test for overall effect: Z = 3.01 (P = 0.003)

Favours (photon) Favours (proton)

(a)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, fixed, 95% CI	Mean difference IV, fixed, 95% CI
Correia, 2019	8.88	2.4	11	1.04	0.075	11	98.1%	7.84 [6.42, 9.26]	
Dennis, 2013	18.66	15.45	11	4.36	7.85	11	1.9%	14.30 [4.06, 24.54]	
Total (95% CI)			22			22	100.0%	7.96 [6.56, 9.37]	

Heterogeneity: chi^2 = 1.50; df = 1 (P = 0.22); I^2 = 33%
Test for overall effect: Z = 11.10 (P < 0.00001)

Favours (photon) Favours (proton)

(b)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, fixed, 95% CI	Mean difference IV, fixed, 95% CI
Correia, 2019	5.4	0.329	11	0.12	0.001	11	100.0%	5.28 [5.09, 5.47]	
Dennis, 2013	20.091	17.423	11	6.273	13.859	11	0.0%	13.82 [0.66, 26.97]	
Total (95% CI)			22			22	100.0%	5.28 [5.09, 5.48]	

Heterogeneity: chi^2 = 1.62; df = 1 (P = 0.20); I^2 = 38%
Test for overall effect: Z = 53.25 (P < 0.00001)

Favours (photon) Favours (proton)

(c)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, fixed, 95% CI	Mean difference IV, fixed, 95% CI
Correia, 2019	4.36	0.209	11	0.16	0.002	11	99.9%	4.20 [4.08, 4.32]	
Stoker, 2018	14.6	3.9	10	8.3	3.5	10	0.1%	6.30 [3.05, 9.55]	
Total (95% CI)			21			21	100.0%	4.20 [4.08, 4.33]	

Heterogeneity: chi^2 = 1.60; df = 1 (P = 0.21); I^2 = 38%
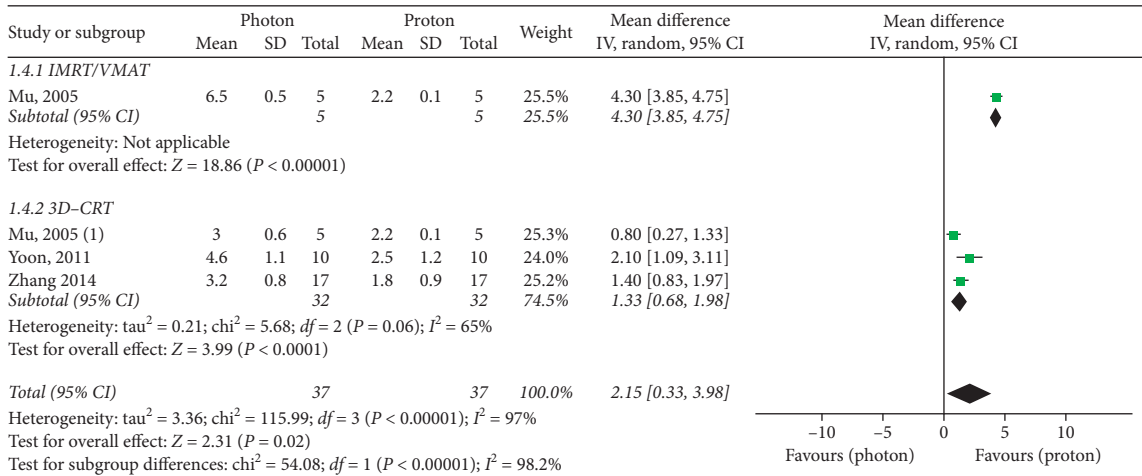Test for overall effect: Z = 66.74 (P < 0.00001)

Favours (photon) Favours (proton)

(d)

Figure 5: Continued.

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI	Mean difference IV, random, 95% CI
1.1.1 IMRT/VMAT									
Mu, 2005	13.2	0.7	5	2.3	0.4	5	36.2%	10.90 [10.19, 11.61]	
Subtotal (95% CI)			5			5	36.2%	*10.90 [10.19, 11.61]*	
Heterogeneity: Not applicable									
Test for overall effect: Z = 30.23 (P < 0.00001)									
1.1.2 3D–CRT									
Mu, 2005 (1)	18.2	0.6	5	2.3	0.4	5	36.3%	15.90 [15.27, 16.53]	
Yoon, 2011	34.6	1.1	10	19.4	6.2	10	27.4%	15.20 [11.30, 19.10]	
Subtotal (95% CI)			15			15	63.8%	*15.88 [15.26, 16.51]*	
Heterogeneity: tau^2 = 0.00; chi^2 = 0.12; df = 1 (P = 0.73); I^2 = 0%									
Test for overall effect: Z = 49.89 (P < 0.00001)									
Total (95% CI)			20			20	100.0%	13.90 [9.82, 17.97]	
Heterogeneity: tau^2 = 11.82; chi^2 = 107.41; df = 2 (P < 0.00001); I^2 = 98%									
Test for overall effect: Z = 6.68 (P < 0.00001)									
Test for subgroup differences: chi^2 = 107.29; df = 1 (P < 0.00001); I^2 = 99.1%									

−100 −50 0 50 100
Favours (photon) Favours (proton)

(e)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI	Mean difference IV, random, 95% CI
Yoon, 2011	29.9	2.1	10	4.1	4.7	10	49.2%	25.80 [22.61, 28.99]	
Zhang, 2014	15.3	3.7	17	0.1	0.1	17	50.8%	15.20 [13.44, 16.96]	
Total (95% CI)			27			27	100.0%	20.41 [10.03, 30.80]	
Heterogeneity: tau^2 = 54.45; chi^2 = 32.51; df = 1 (P < 0.00001); I^2 = 97%									
Test for overall effect: Z = 3.85 (P = 0.0001)									

−100 −50 0 50 100
Favours (photon) Favours (proton)

(f)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI	Mean difference IV, random, 95% CI
1.4.1 IMRT/VMAT									
Mu, 2005	6.5	0.5	5	2.2	0.1	5	25.5%	4.30 [3.85, 4.75]	
Subtotal (95% CI)			5			5	25.5%	*4.30 [3.85, 4.75]*	
Heterogeneity: Not applicable									
Test for overall effect: Z = 18.86 (P < 0.00001)									
1.4.2 3D–CRT									
Mu, 2005 (1)	3	0.6	5	2.2	0.1	5	25.3%	0.80 [0.27, 1.33]	
Yoon, 2011	4.6	1.1	10	2.5	1.2	10	24.0%	2.10 [1.09, 3.11]	
Zhang 2014	3.2	0.8	17	1.8	0.9	17	25.2%	1.40 [0.83, 1.97]	
Subtotal (95% CI)			32			32	74.5%	*1.33 [0.68, 1.98]*	
Heterogeneity: tau^2 = 0.21; chi^2 = 5.68; df = 2 (P = 0.06); I^2 = 65%									
Test for overall effect: Z = 3.99 (P < 0.0001)									
Total (95% CI)			37			37	100.0%	2.15 [0.33, 3.98]	
Heterogeneity: tau^2 = 3.36; chi^2 = 115.99; df = 3 (P < 0.00001); I^2 = 97%									
Test for overall effect: Z = 2.31 (P = 0.02)									
Test for subgroup differences: chi^2 = 54.08; df = 1 (P < 0.00001); I^2 = 98.2%									

−10 −5 0 5 10
Favours (photon) Favours (proton)

(g)

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI	Mean difference IV, random, 95% CI
Yoon, 2011	8	1.3	10	0.3	0.2	10	49.1%	7.70 [6.88, 8.52]	
Zhang, 2014	5.1	0.7	17	0.2	0.1	17	50.9%	4.90 [4.56, 5.24]	
Total (95% CI)			27			27	100.0%	6.27 [3.53, 9.02]	
Heterogeneity: tau^2 = 3.82; chi^2 = 38.73; df = 1 (P < 0.00001); I^2 = 97%									
Test for overall effect: Z = 4.48 (P < 0.00001)									

−20 −10 0 10 20
Favours (photon) Favours (proton)

(h)

FIGURE 5: Continued.

Study or subgroup	Photon Mean	SD	Total	Proton Mean	SD	Total	Weight	Mean difference IV, random, 95% CI	Mean difference IV, random, 95% CI
1.6.1 IMRT/VMAT									
Mu, 2005	8.9	0.7	5	1.3	0.1	5	33.6%	7.60 [6.98, 8.22]	
Subtotal (95% CI)			5			5	33.6%	*7.60 [6.98, 8.22]*	
Heterogeneity: not applicable									
Test for overall effect: Z = 24.03 (P < 0.00001)									
1.6.2 3D–CRT									
Mu, 2005 (1)	2.7	0.5	5	1.3	0.1	5	33.7%	1.40 [0.95, 1.85]	
Yoon, 2011	4.3	1.5	10	2.2	1.6	10	32.8%	2.10 [0.74, 3.46]	
Subtotal (95% CI)			15			15	66.4%	*1.47 [1.04, 1.89]*	
Heterogeneity: tau^2 = 0.00; chi^2 = 0.92; df = 1 (P = 0.34); I^2 = 0%									
Test for overall effect: Z = 6.78 (P < 0.00001)									
Total (95% CI)			20			20	100.0%	3.71 [–0.81, 8.23]	
Heterogeneity: tau^2 = 15.76; chi^2 = 256.81; df = 2 (P < 0.00001); I^2 = 99%									
Test for overall effect: Z = 1.61 (P = 0.11)									
Test for subgroup differences: chi^2 = 255.89; df = 1 (P < 0.00001); I^2 = 99.6%									

–20 –10 0 10 20
Favours (photon)   Favours (proton)

(i)

FIGURE 5: Dmean and Dmax of extracranial OARs. (a) Dmean - left lens. (b) Dmean - left lacrimal gland. (c) Dmean - right lacrimal gland. (d) Dmax - left lens. (e) Dmean - esophagus. (f) Dmean - thyroid. (g) Dmean - lungs. (h) Dmean - liver. (i) Dmean - kidneys. Mu, 2005: IMRT; Mu, 2005 (1): 3D-CRT.

(WV) RT (24–30 Gy) for patients affected by germ-cell tumors, Stoker et al. [14] analyzed hippocampal-avoidance whole-brain RT (total dose: 36 Gy), while other authors analyzed RT to the resected intracranial tumor bed with total target doses in the range between 50.4 and 55.8 Gy (Table 2).

Because of the limited number of included studies, we did not perform subgroup analyses for OARs according to the extent of treatment fields (extended/localized) or total target doses (low/high), but we observed the advantages of PBT in any cases of OAR dose sparing—also including the in-field organs (e.g., brainstem and normal brain) and considering the comparison with modern (high-conformal) photon RT techniques (IMRT/VMAT).

Furthermore, the influence of the OAR's location on absorbed doses can be easily understood: also Howell [13] considered that, during CSI, lungs and kidneys are located bilaterally to the target volume, so they received higher doses in PBT plans as compared to anterior OARs such as the esophagus. This difference observed in spinal treatments is mainly due to the physical properties of the proton beam. Our subgroup analyses according to photon techniques showed a better dose sparing (higher MD) for lateral organs (such as the lungs and kidneys) when PBT is compared to intensity-modulated techniques, and an improved dose sparing of anterior organs (such as the esophagus) when PBT is compared to 3D-CRT. These findings could be related to different beam arrangement between the photon RT techniques.

Globally, the dosimetric benefits for the *in-field*, *partially in-field*, and *out-of-field* OARs obtained with PBT [13]—which are due to the characteristic dose distribution of PBT, dependent on physical properties of protons [13]—could translate into the reduction of neurocognitive damages, visual and hearing loss, endocrine dysfunction, and other late toxicities and SMNs risk. Higher-level evidences from appropriate studies and long-term clinical data are still needed to confirm these suggestions.

With particular regard for SMNs risk assessment, even if we did not specifically analyze risk models, we observed that secondary cancer risks were assessed in some included studies [15, 17–19] that suggested a probabilistic benefit with PBT. Nevertheless, several authors [13, 15, 16] emphasized the concern of neutron contamination risk related to proton treatments in the assessment of SMN risk. On this topic, Beltran, however, observed that a quite low neutron dose for IMPT was reported [16, 33]. More recently, Schneider and Halg [34] have underlined the limitations of previous risk models which assessed the impact of neutron dose. The authors [34] suggested a reduction of SMN risk with PBT when adequate risk models—that take into account well-calculated dose distributions—are used. Globally, it has to be noted that thanks to a reduced integral dose, the risk of SMNs remains lower for PBT as compared to that for photon RT [16, 35], as observed in studies on pediatric CSI that took into account neutron contamination [10, 36] and in a wide retrospective analysis including adult patients [37].

4.4. Study Limitations and Additional Considerations. Despite recent advances in radiobiological knowledge, the evaluation of RISEs risk in pediatric patients is still difficult because of the particular radiation sensitivity of developing tissues [1, 38] and the lack of comprehensive radiation dose-volume data in this setting [38]. Indeed, the most used dose constraints for normal tissues reported by QUANTEC (Quantitative Analysis of Normal Tissue Effects in the Clinic) [39] are referred to adults [38] treated with photons. A review of dose constraints and recommendations for intracranial organs at risk (OARs) for both adult and pediatric patients was published in 2015 by Scoccianti et al. [27], even if pediatric constraints were reported in few studies. The choice to consider the Dmean and Dmax values as referring parameters for our secondary analysis was based on an overview of reported normal tissues dose constraints in pediatrics [27–29]. Nevertheless, we are conscious that more informative data are required, such as those which are being expected by collaborative long-term observational
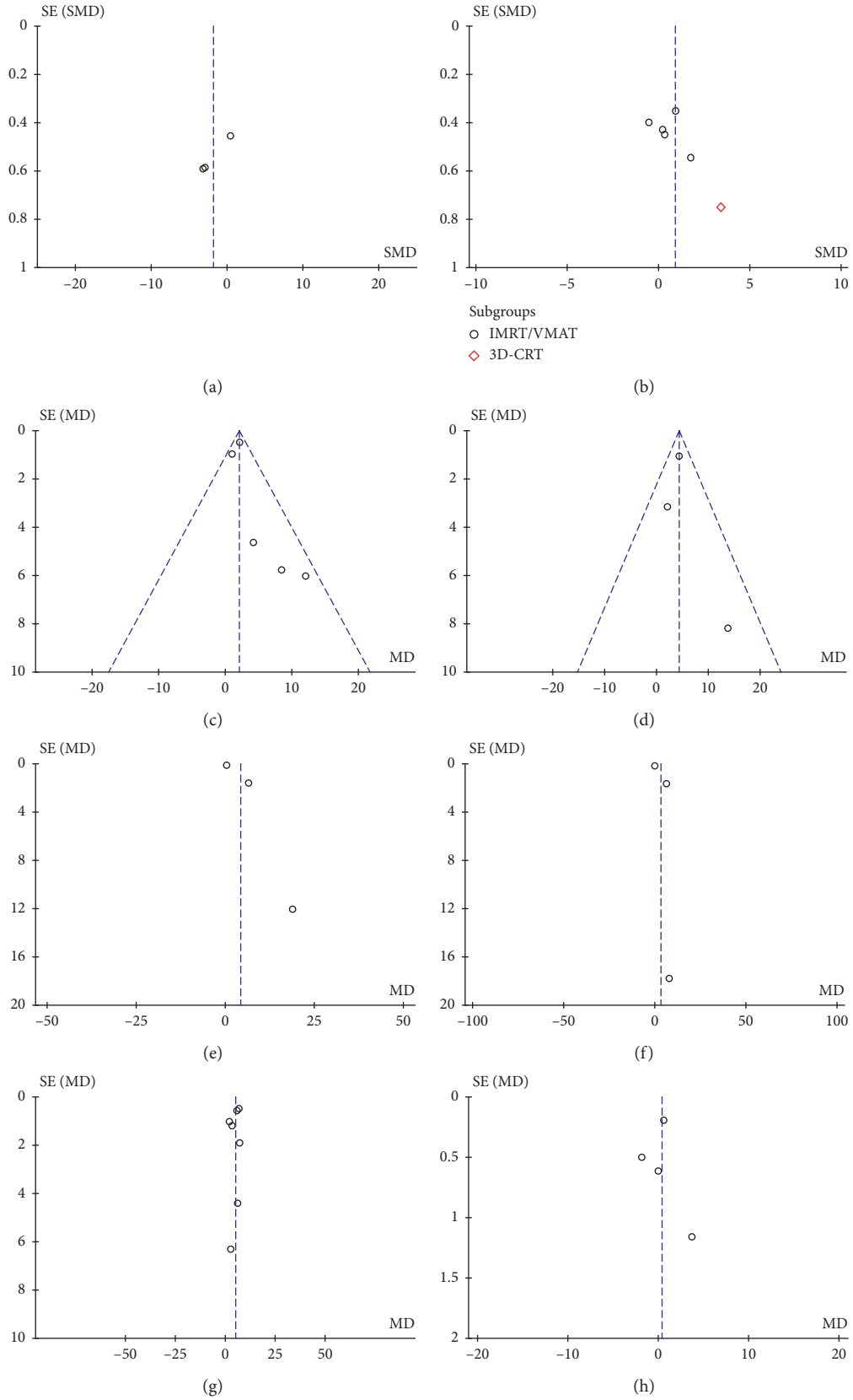
FIGURE 6: Funnel plots for target and intracranial OARs. (a) Conformity Index. (b) Homogeneity Index. (c) Dmean of the brainstem. (d) Dmean of the optic chiasm. (e) Dmean of the left cochlea. (f) Dmean of the right cochlea. (g) Dmean of the normal brain. (h) Dmax of the normal brain.
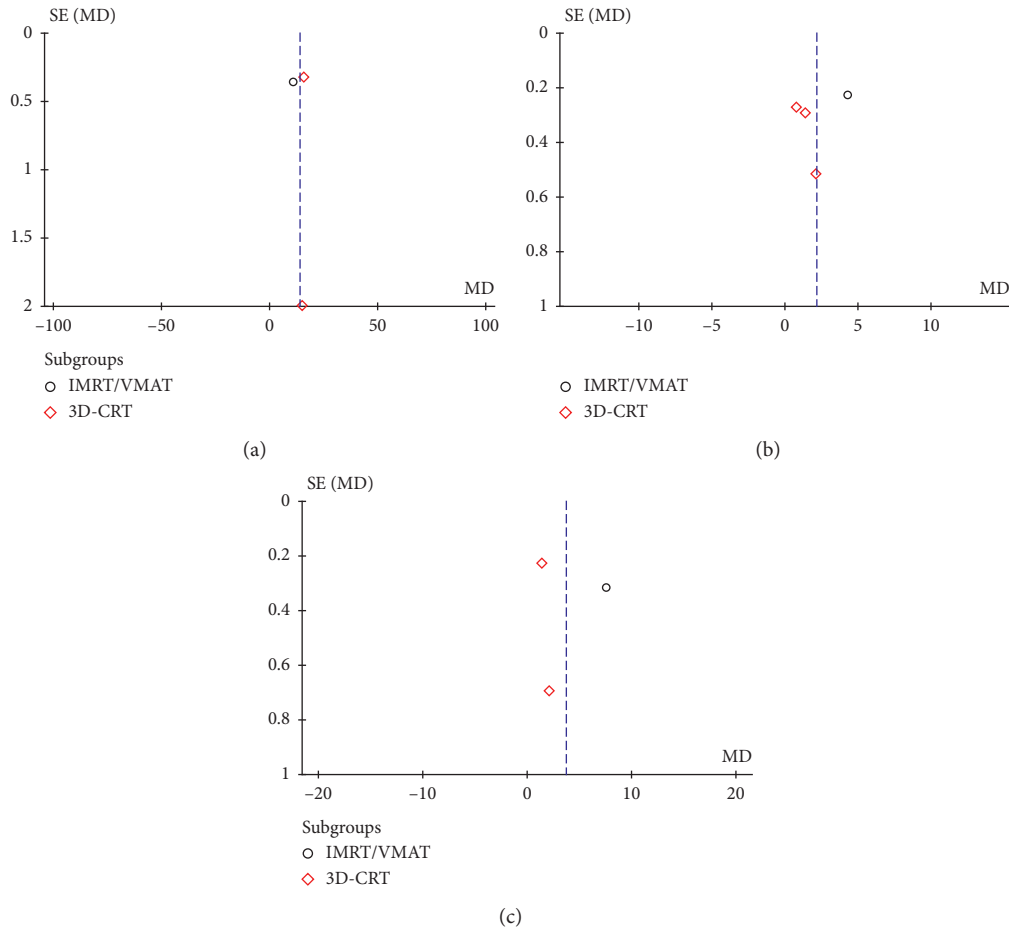
FIGURE 7: Funnel plots for extracranial OARs. (a) Dmean of the esophagus. (b) Dmean of the lungs. (c) Dmean of the kidneys.

studies and ongoing clinical research programs (see the PENTEC (Pediatric Normal Tissue Effects in the Clinic) group project [38]).

The lack of data collected from randomized-controlled trials and the absence of a risk of bias assessment for the individual studies could be considered as limitations of our analysis; nevertheless, we remark that the primary purpose of our work was a secondary analysis of dosimetric comparison studies, which was anyhow feasible because all the included studies analyzed the same patient cohort and compared plans generated with the same treatment planning software.

We however observed a major study limitation in the lack of useful dosimetric data for all the considered OARs: this made us unable to perform a comprehensive secondary analysis for all cases (e.g., when analyzed dosimetric data were reported by single studies) or a complete evaluation of publication bias. Informative data could also have been lost because of specific requirements of our research strategy (e.g., studies reporting the range of average mean doses instead of SD, as well as different dose/volume constraints that were excluded). Also, the limited number and the heterogeneity (heterogeneity can be related to factors such as sex, age, height, and weight that could influence morphometric profiles) of enrolled patients in the included studies can be considered as study limitations.

Furthermore, publication bias was observed for Conformity Index analysis. Indeed, the heterogeneity in the calculation of the considered target dosimetric parameters—which is due to the absence of univocal formulae—could have introduced a potential limitation in our secondary analyses, particularly for Conformity Index assessment.

To reduce the risk of inconsistent results in Conformity Index assessment, we chose to analyze studies [10, 12, 16] that provided a comparison according to analogous formulae [24]. We also agreed with the observation by Ho et al. [6] on the inappropriateness of Conformity Index as the referring parameter for target dose conformation when large target volumes—as those in CSI—are considered. For all these reasons, we excluded Yoon et al. [15] and Howell et al.'s [13] studies from our analysis.

We however considered that a reasonably high level of concordance between different formulae for the calculation of Homogeneity Index has been demonstrated [25]. Additionally, to reduce the potential heterogeneity among studies that assessed target homogeneity, we performed subgroup analyses according to photon RT techniques characterized by different conformation properties. Finally, we remark that we used the standardized mean difference (SMD)—according to the Cochrane recommendations (http://handbook-5-1.cochrane.org)—as a summary statistic to

take into account studies that assessed the same outcomes but measured them using a variety of formulae.

Because of the limitations of this meta-analysis, we suggest that the reported results have to be correlated with long-term follow-up data from well-designed studies with larger samples to provide significant information useful in clinical practice. This goal could be more easily achieved, thanks to comprehensive database, as suggested also by Weber et al. [22]. Indeed, the realization of modeling studies for more accurate dose-response and toxicity assessments could benefit from data sharing.

Lastly, we disclose we did not analyze specific dosimetric data for vertebral structures. Only in 2019, a consensus has been published for dose constraints for these structures [40]. Because of the relevance of vertebral bone exposure to children growth, an overview of previously published dosimetric comparison studies—taking into account current dose recommendations—is encouraged to better assess the potential of PBT.

5. Conclusions

Our analysis supports current knowledge concerning the dosimetric advantages of PBT over photon RT for pediatric CNS tumors. Protons improve the dose sparing of in-field and out-of-field OARs located in both intracranial and extracranial districts while maintaining satisfactory target conformity and homogeneity. These dosimetric advantages could lead to clinical improvements in pediatric radiation treatments. Wider dosimetric data are necessary to improve the quality of evidence, and further clinical studies and cost-effective analyses comparing photon and proton treatments are required to confirm the benefits of PBT in clinical practice.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Supplementary Materials

Table: results of single studies reporting dosimetric comparisons between photon RT and PBT for intracranial and extracranial OARs. (*Supplementary Materials*)

References

[1] T. Steinmeier, S. Schulze Schleithoff, and B. Timmermann, "Evolving radiotherapy techniques in paediatric oncology," *Clinical Oncology*, vol. 31, no. 3, Article ID 142e150, 2019.

[2] G. T. Armstrong, Q. Liu, Y. Yasui et al., "Long-term outcomes among adult survivors of childhood central nervous system malignancies in the childhood cancer survivor Study," *JNCI: Journal of the National Cancer Institute*, vol. 101, no. 13, pp. 946–958, 2009.

[3] M. Mizumoto, Y. Oshiro, T. Yamamoto, H. Kohzuki, and H. Sakurai, "Proton beam therapy for pediatric brain tumor," *Neurologia Medico-Chirurgica*, vol. 57, no. 7, pp. 343–355, 2017.

[4] A. Laprie, Y. Hu, C. Alapetite et al., "Paediatric brain tumours: a review of radiotherapy, state of the art and challenges for the

future regarding protontherapy and carbontherapy," *Cancer/Radiothérapie*, vol. 19, no. 8, pp. 775–789, 2015.

[5] E. Ludmir, K. Grosshans, and K. D. Woodhouse, "Radiotherapy advances in pediatric neuro-oncology," *Bioengineering*, vol. 5, no. 4, p. 97, 2018.

[6] E. S. Q. Ho, S. A. Barrett, and L. M. Mullaney, "A review of dosimetric and toxicity modeling of proton versus photon craniospinal irradiation for pediatrics medulloblastoma," *Acta Oncologica*, vol. 56, no. 8, pp. 1031–1042, 2017.

[7] R. Leroy, N. Benahmed, F. Hulstaert, N. Van Damme, and D. De Ruysscher, "Proton therapy in children: a systematic review of clinical effectiveness in 15 pediatric cancers," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 95, no. 1, pp. 267–278, 2016.

[8] C. Main, M. Dandapani, M. Pritchard et al., "The effectiveness and safety of proton beam radiation therapy in children with malignant central nervous system (CNS) tumours: protocol for a systematic review," *Systematic Reviews*, vol. 5, no. 1, p. 124, 2016.

[9] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement," *PLoS Medicine*, vol. 6, no. 7, Article ID e1000097, 2009.

[10] N. S. Boehling, D. R. Grosshans, J. B. Bluett et al., "Dosimetric comparison of three-dimensional conformal proton radiotherapy, intensity-modulated proton therapy, and intensity-modulated radiotherapy for treatment of pediatric craniopharyngiomas," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 82, no. 2, pp. 643–652, 2012.

[11] D. Correia, D. Terribilini, S. Zepter et al., "Whole-ventricular irradiation for intracranial germ cell tumors: dosimetric comparison of pencil beam scanned protons, intensity-modulated radiotherapy and volumetric-modulated arc therapy," *Clinical and Translational Radiation Oncology*, vol. 15, pp. 53–61, 2019.

[12] D. Freund, R. Zhang, M. Sanders, and W. Newhauser, "Predictive risk of radiation induced cerebral necrosis in pediatric brain cancer patients after VMAT versus proton therapy," *Cancers*, vol. 7, no. 2, pp. 617–630, 2015.

[13] R. M. Howell, A. Giebeler, W. Koontz-Raisig et al., "Comparison of therapeutic dosimetric data from passively scattered proton and photon craniospinal irradiations for medulloblastoma," *Radiation Oncology*, vol. 7, no. 1, p. 116, 2012.

[14] J. Stoker, S. Vora, A. Patel et al., "Advantages of intensity modulated proton therapy during hippocampal avoidance whole brain radiation therapy," *Physics and Imaging in Radiation Oncology*, vol. 8, pp. 28–32, 2018.

[15] M. Yoon, D. H. Shin, J. Kim et al., "Craniospinal irradiation techniques: a dosimetric comparison of proton beams with standard and advanced photon radiotherapy," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 81, no. 3, pp. 637–646, 2011.

[16] C. Beltran, M. Roca, and T. E. Merchant, "On the benefits and risks of proton therapy in pediatric craniopharyngioma," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 82, no. 2, pp. e281–e287, 2012.

[17] E. R. Dennis, M. R. Bussière, A. Niemierko et al., "A comparison of critical structure dose and toxicity risks in patients with low grade gliomas treated with IMRT versus proton radiation therapy," *Technology in Cancer Research & Treatment*, vol. 12, no. 1, pp. 1–9, 2013.

[18] X. Mu, T. Björk-Eriksson, S. Nill et al., "Does electron and proton therapy reduce the risk of radiation induced cancer

after spinal irradiation for childhood medulloblastoma? A comparative treatment planning study," *Acta Oncologica*, vol. 44, no. 6, pp. 554–562, 2005.

[19] R. Zhang, R. M. Howell, P. J. Taddei, A. Giebeler, A. Mahajan, and W. D. Newhauser, "A comparative study on the risks of radiogenic second cancers and cardiac mortality in a set of pediatric medulloblastoma patients treated with photon or proton craniospinal irradiation," *Radiotherapy and Oncology*, vol. 113, no. 1, pp. 84–88, 2014.

[20] D. N. Louis, A. Perry, G. Reifenberger et al., "The 2016 World Health Organization classification of tumors of the central nervous system: a summary," *Acta Neuropathologica*, vol. 131, no. 6, pp. 803–820, 2016.

[21] L. Padovani, G. Horan, and T. Ajithkumar, "Radiotherapy advances in paediatric medulloblastoma treatment," *Clinical Oncology*, vol. 31, no. 3, Article ID 171e181, 2019.

[22] D. C. Weber, J. L. Habrand, B. S. Hoppe et al., "Proton therapy for pediatric malignancies: fact, figures and costs. A joint consensus statement from the pediatric subcommittee of PTCOG, PROS and EPTN," *Radiotherapy and Oncology*, vol. 128, no. 1, pp. 44–55, 2018.

[23] E. Seravalli, M. Bosman, Y. Lassen-Ramshad et al., "Dosimetric comparison of five different techniques for craniospinal irradiation across 15 European centers: analysis on behalf of the SIOP-E-BTG (radiotherapy working group)," *Acta Oncologica*, vol. 57, no. 9, pp. 1240–1249, 2018.

[24] K. Ohtakara, S. Hayashi, H. Tanaka, and H. Hoshi, "Consideration of optimal isodose surface selection for target coverage in micro-multileaf collimator-based stereotactic radiotherapy for large cystic brain metastases: comparison of 90%, 80% and 70% isodose surface-based planning," *The British Journal of Radiology*, vol. 85, no. 1017, pp. e640–e646, 2012.

[25] T. Kataria, K. Sharma, V. Subramani, K. Karrthick, and S. Bisht, "Homogeneity index: an objective tool for assessment of conformal radiation treatments," *Journal of Medical Physics*, vol. 37, no. 4, pp. 207–213, 2012.

[26] S. M. MacDonald, S. Safai, A. Trofimov et al., "Proton radiotherapy for childhood ependymoma: initial clinical outcomes and dose comparisons," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 71, no. 4, pp. 979–986, 2008.

[27] S. Scoccianti, B. Detti, D. Gadda et al., "Organs at risk in the brain and their dose-constraints in adults and in children: a radiation oncologist's guide for delineation in everyday practice," *Radiotherapy and Oncology*, vol. 114, no. 2, pp. 230–238, 2015.

[28] I. Kristensen, P. Nilsson, and N. Nilsson, "Comparative proton and photon treatment planning in pediatric patients with various diagnoses," *International Journal of Particle Therapy*, vol. 2, no. 2, pp. 367–375, 2015.

[29] N. P. Brodin, P. M. A. Rosenschöld, M. C. Aznar et al., "Radiobiological risk estimates of adverse events and secondary cancer for proton and photon radiation therapy of pediatric medulloblastoma," *Acta Oncologica*, vol. 50, no. 6, pp. 806–816, 2011.

[30] B. R. Eaton and T. Yock, "The use of proton therapy in the treatment of benign or low-grade pediatric brain tumors," *The Cancer Journal*, vol. 20, no. 6, pp. 403–408, 2014.

[31] D. Takizawa, M. Mizumoto, T. Yamamoto et al., "A comparative study of dose distribution of PBT, 3D-CRT and IMRT for pediatric brain tumors," *Radiation Oncology*, vol. 12, no. 1, p. 40, 2017.

[32] T. E. Merchant, E. N. Kiehna, C. Li et al., "Modeling radiation dosimetry to predict cognitive outcomes in pediatric patients with CNS embryonal tumors including medulloblastoma," *International Journal of Radiation Oncology, Biology, Physics*, vol. 65, pp. 210–212, 2006.

[33] U. Schneider, S. Agosteo, E. Pedroni, and J. Besserer, "Secondary neutron dose during proton therapy using spot scanning," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 53, no. 1, pp. 244–251, 2002.

[34] W. Schneider and R. Hälg, "Impact of neutron in clinical proton therapy," *Frontiers in Oncology*, vol. 5, p. 235, 2015.

[35] R. Miralbell, A. Lomax, L. Cella, and U. Schneider, "Potential reduction of the incidence of radiation-induced second cancers by using proton beams in the treatment of pediatric tumors," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 54, no. 3, pp. 824–829, 2002.

[36] W. D. Newhauser, J. D. Fontenot, A. Mahajan et al., "The risk of developing a second cancer after receiving craniospinal proton irradiation," *Physics in Medicine and Biology*, vol. 54, no. 8, pp. 2277–2291, 2009.

[37] C. S. Chung, T. I. Yock, K. Nelson, Y. Xu, N. L. Keating, and N. J. Tarbell, "Incidence of second malignancies among patients treated with proton versus photon radiation," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 87, no. 1, pp. 46–52, 2013.

[38] L. S. Constine, C. M. Ronckers, C.-H. Hua et al., "Pediatric normal tissue effects in the clinic (PENTEC): an international collaboration to analyse normal tissue radiation dose-volume response relationships for paediatric cancer patients," *Clinical Oncology*, vol. 31, no. 3, pp. 199–207, 2019.

[39] L. B. Marks, E. D. Yorke, A. Jackson et al., "Use of normal tissue complication probability models in the clinic," *International Journal of Radiation Oncology∗Biology∗Physics*, vol. 76, no. 3, pp. S10–S19, 2010.

[40] B. A. Hoeben, C. Carrie, B. Timmermann et al., "Management of vertebral radiotherapy dose in paediatric patients with cancer: consensus recommendations from the SIOPE radiotherapy working group," *Lancet Oncology*, vol. 20, no. 3, 2019.