Automated Extraction of Dose/Volume Statistics for Radiotherapy-Treatment-Plan Evaluation in Clinical-Trial Quality Assurance

Yutao U. T. Gong[1], Jialu Yu[1,2]*, Dalong Pang[2], Heming Zhen[3], James Galvin[1] and Ying Xiao[1,4]

[1] IROC, Philadelphia, PA, USA, [2] Department of Radiation Medicine, Georgetown University Hospital, Washington, DC, USA, [3] Rush University Medical Center, Chicago, IL, USA, [4] Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA, USA

Radiotherapy clinical-trial quality assurance is a crucial yet challenging process. This note presents a tool that automatically extracts dose/volume statistics for determining dosimetry compliance review with improved efficiency and accuracy. A major objective of this study is to develop an automated solution for clinical-trial radiotherapy dosimetry review.

Keywords: radiation therapy plan evaluation, clinical trial, quality assurance, automated data processing, Matlab

INTRODUCTION

Quality assurance (QA) of radiation therapy (RT) procedures that include treatment-plan review is an important mechanism to ensure adequate dose coverage for tumor(s), while identifying risks of normal tissue complications (1–3). The Imaging and Radiation Oncology Core (IROC) provides RT QA services to the National Clinical Trial Network (NCTN). IROC services include site qualification, trial-design support, credentialing, data management, and case review (4–6). Case review includes a data-integrity review, a review of compliance with target volume and organs at risk contours by study chairs, and a review of dosimetry compliance (5). The dosimetry review (4, 7–11) checks whether the data points on planned structure dose–volume histograms (DVH) meet the protocol criteria. Non-compliance of target volume–dose criteria may result in treatment failure, and violation of organs at risk criteria may create complications and toxicities. It is a consensus that timely clinical-trial data QA is crucial for the successful RT clinical-trial management (4, 7, 8, 10, 12–15). A study of 174 medulloblastoma patients in 1999 found that the number of major target deviations in RT was strongly correlated with the risk of tumor relapse (16). Another study of 416 pancreatic cancer patients found that failure to adhere to protocol-specified radiation-therapy guidelines was associated with decreased survival (17). One published meta-analysis about radiotherapy-protocol deviations and clinical outcomes (18) included two lung cancer trials (19, 20), three trials for medulloblastoma or supratentorial primitive neuroectodermal tumors (16, 21, 22), and one trial each for Ewing sarcoma (23), pancreatic cancer (17), and head and neck cancer (7). The study found the frequency of RT QA deviations ranged from 8 to 71% (median 32%), and RT deviations were associated with a statistically significant decrease in overall survival and also secondary outcomes (18).

Automatic data processing is needed to improve the efficiency of RT QA, which at present is a crucial yet labor-intensive and challenging task (4). Uniformity in the data format used for information collection is a prerequisite of automatic data processing. Recently, infrastructures and guidelines were developed to improve the data uniformity within NCTN (24, 25). A submission-ready case should have all the structures required by the protocol, and the structure names should exactly follow the standard names as stated in each protocol (24). The dosimetry compliance review usually is a task of checking more than 10-dose/volume points for agreement with stated protocol limits (**Table 1**). Considering that typically 100 or more patients are to be recruited per trial and there are more than 50 RT open trials on the NRG Oncology website (26), a tool that automatically extracts dose–volume points for dosimetry compliance review will save time, improve accuracy, and catch underlying issues with understanding of the protocol text. In this report, we introduce an automated solution that meets the need for rapid dosimetry review and data extraction.

METHODS

Workflow of the Solution
The proposed automated solution to QA for clinical-trial data management consists of three separate parts: (i) a standard syntax system that defines the plan–evaluation data points, (ii) an automated program with graphic user interface (GUI) written in MATLAB, and (iii) a result output MATLAB program that sorts and lists the results as a spreadsheet. The combination of the MATLAB programs is called dose/volume point statistics (DPS) (27). The workflow of this solution is illustrated in **Figure 1**. DPS takes user input (structure names and expressions of dose–volume points) and DVH files from DVH reviewer of MIM (MIM Software Inc., Cleveland, OH, USA) and automatically completes computation and statistics, and then generates reports. The DVH files referred to in this report use the comma-separated value format and contain more information than the simple DVH values (see Format of DVH File for more detailed information).

The GUI for DPS is shown in **Figure 2**. The data extraction requests are specified by entering the structure names and names of dose/volume points in DPS. The request can be exported and imported as a plain text file. The run button in GUI loads the DVH files. Details for preparing the structure names, naming of both dose/volume points and DVH files can be found in Sections "Structure Names," "Standard Syntax System," and "Format of

DVH File." For calculating the normalized dose (e.g., the ratio of max dose to the prescription dose), a global dose reference point is set. DPS allows two choices of the dose reference point: the prescription dose and the maximum dose. The prescription dose can be entered in DPS as a global parameter for all cases; however, the maximum dose must be from the individual DVH files. DPS is designed to accommodate protocols with variable prescription doses. A specific prescription dose value saved in the DVH file overwrites the globally defined prescription dose.

In operation, DPS extracts DVH data of the structure in the first row of GUI from the first DVH files, and it calculates values of the dose/volume points in the same row of GUI. Then, DPS moves to the next row of GUI until all dosimetry-review requests are processed for the case and moves on to the next case. When all cases are processed, DPS tabulates the dose/volume points for all cases in one spreadsheet report and calculates 1, 2, 5, 10, 15, 85, 90, 95, 98, and 99% quantiles of each dose/volume point (**Figure 3**). This report can be used to evaluate the distribution of plan qualities for a particular protocol. Based on this information, it is possible to make the decision to amend a protocol to improve the performance of institutions' enrollments.

Structure Names
Although DPS accepts and processes any structure names defined by users, in order to avoid errors and confusion it is highly recommended to name the structures according to the consensus of the NCTN structure-naming convention defined by Yu et al. (24). While processing a large number of cases, maximum performance of DPS is achievable by sorting case data by different arms of a protocol and running DPS separately for each arm.

Standard Syntax System
Dosimetry-review data points specified in protocols can be entered in DPS using a standard syntax system (**Table 2**). This syntax system intends to avoid possible confusion between absolute values and relative values, as well as the inconvenience of using subscripts. DPS checks syntax on all the input expressions and reminds users of any non-acceptable expressions.
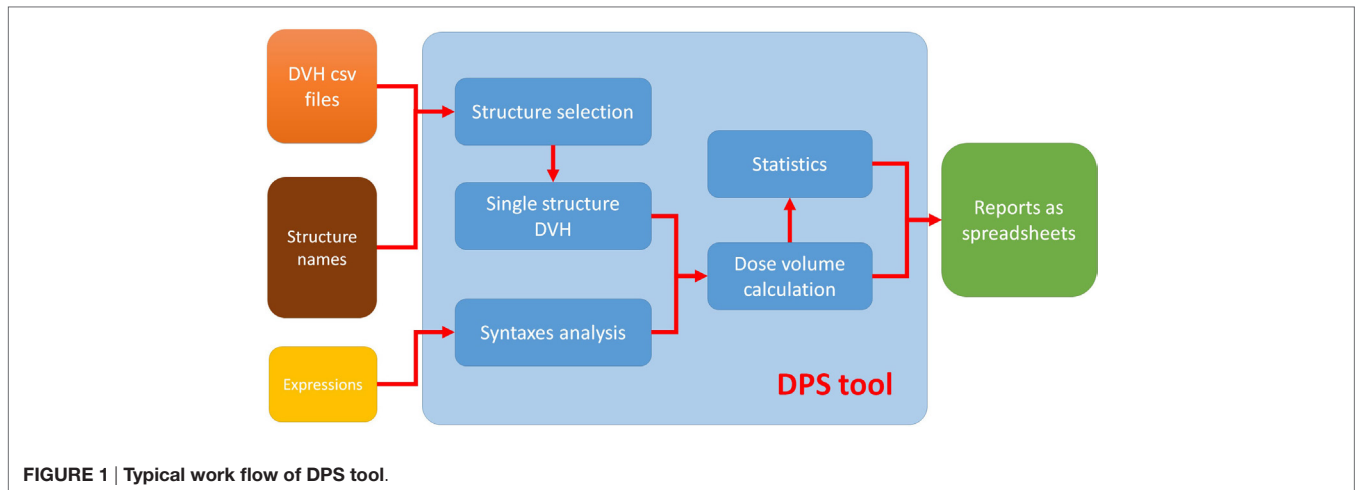
Generally, in the proposed syntax system, "D" stands for dose with the units (Gy or %) defined in the following pair of parentheses, whereas "V" stands for volume with the units (cc or %). **Table 2** describes the syntaxes and their definitions. After setting the dose reference point, which is set to be 100% of the relative dose, the user is able to calculate percentage dose to meet special needs. Popular dose reference points include prescription dose and maximum dose. In addition, simple calculation is supported by the syntax system. For example, Dmean/Rx is the ratio of mean dose to prescription dose.

Format of DVH File
For NRG, all the RT data for submission-ready cases are saved in commercial software for review (MIMvista, Cleveland, OH, USA). The structure-dosimetry data for each case can be exported as a DVH file that is written in a comma-separated value format. These files are sometimes identified as DVH csv files. As shown in **Figure 4**, the first line (blue region) contains the patient/site ID, and the second line (yellow region) has structure names. Numeric

TABLE 1 | Dosimetry-review workload of selected NRG trials (26).

Trial	No. of structures requiring review	No. of data points to be reviewed	Target accrual for trial
HN001	17	18	758
HN002	9	14	296
GU001	3	5	185
LU001	6	14	168
BN001	13	15	576
GI001	10	16	182

FIGURE 1 | **Typical work flow of DPS tool**.

FIGURE 2 | **Panel of dose/volume points statistics (DPS)**.

data are stored at line 3 and below, in which the first column is the dose data (red region), and the rest (green region) contains the absolute volume data of different structures.

As mentioned in Section "Workflow of the Solution," DPS accommodates protocols with variable prescription dose. Case-specific prescription dose can be saved in cell C1 of the DVH file ("Rx = 60 Gy," or in short, "60"). When dose reference point for dosimetry review is set to be the global maximum dose, the maximum dose for each case needs to be saved in cell D1 of the DVH csv file ("Dmax = 80 Gy," or in short, "80"), since different plans usually do not share the same global maximum dose.

Performance of DPS

To test the performance of this tool, data extraction for dosimetry review of 20 DVH csv files was performed using NRG RT trial

LU001 protocol [28]. The time needed to process the task with DPS is recorded. As a comparison of the efficiency, the same task is repeated in a manual way by reading dose/volume points from MIMvista and typing values in a spreadsheet.

RESULTS

To test the performance of this tool, data extraction for dosimetry review of 20 DVH csv files was performed using NRG RT trial LU001 protocol [28]. Including the time for user operation such as selecting DVH files, it took DPS 3 min to generate the final report in Excel with the computer setting as Windows 7 Professional, Intel Xeon CPU 3.20 GHz, and 12.0 GB RAM. The report includes a spreadsheet for each structure as well as an overview spreadsheet for all the structures (**Figure 3**). The

	A	B	C	D	E	F	G	H	I
1		PTV_6000	PTV_6000	PTV_6000	SpinalCord	Lungs	Lungs	Lungs	Esophagus
2	case_ID	D(Gy)95(%)	Dmin	D(Gy)0.03(cc)	D(Gy)0.03(cc)	V(%)5(Gy)	V(%)20(Gy)	Dmean	V(%)35(Gy)
3	sample_01	59.32	55.43	73.53	46.21	57.50	31.11	14.36	32.22
4	sample_02	60.59	58.85	71.22	30.59	49.48	24.47	22.93	36.82
5	sample_03	59.62	55.02	68.51	24.67	63.83	35.29	16.02	50.83
6	sample_04	60.06	58.10	70.13	31.16	46.05	28.32	12.59	30.76
7	sample_05	59.33	58.27	69.02	18.86	61.98	23.77	22.67	54.15
8	sample_06	60.20	58.47	65.76	20.28	45.87	38.48	23.72	48.99
				. . . . . .					
17	sample_15	59.30	55.58	69.91	16.72	54.30	36.11	13.10	42.71
18	sample_16	60.65	55.54	69.89	21.76	49.80	19.87	11.64	46.23
19	sample_17	60.08	58.48	68.38	40.96	56.27	20.65	14.15	47.66
20	sample_18	60.99	57.32	74.00	44.27	56.34	19.27	14.46	40.28
21	sample_19	59.16	57.20	68.69	40.91	66.16	20.69	15.94	39.55
22	sample_20	59.89	55.58	66.11	33.04	65.43	25.89	17.11	55.69
23	1%	59.16	55.02	65.50	15.62	42.60	19.27	11.56	30.76
24	2%	59.16	55.02	65.50	15.62	42.60	19.27	11.56	30.76
25	5%	59.16	55.18	65.63	16.17	44.23	19.57	11.60	31.49
26	10%	59.24	55.38	65.94	17.06	45.96	20.03	12.12	34.12
27	15%	59.31	55.49	66.17	18.12	47.76	20.42	12.85	36.10
28	85%	60.62	58.48	73.77	47.03	65.20	35.70	21.31	52.49
29	90%	60.74	58.56	74.01	50.26	65.80	36.50	22.80	54.60
30	95%	60.91	58.75	74.24	52.96	67.95	37.69	23.32	55.36
31	98%	60.99	58.85	74.45	53.25	69.74	38.48	23.72	55.69
32	99%	60.99	58.85	74.45	53.25	69.74	38.48	23.72	55.69

FIGURE 3 | **Sample report spreadsheet for LU001 opened in Excel**. The blue region on the top shows dosimetry-review requests entered in GUI, and the green region contains values of dose–volume points calculated by DPS. The red region is the statistics of values in the green region. Data shown in the table are made examples and are not from submitted cases.

	A	B	C	D	E	F	G
1		Patient/Site ID	Rx=60Gy	Dmax=80Gy			
2		PTV	Larynx	Parotid_R	Parotid_L	Eye_L	Eye_R
3	0	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
4	0.1	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
5	0.2	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
6	0.3	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
7	0.4	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
8	0.5	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
9	0.6	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
10	0.7	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
11	0.8	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
12	0.9	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
13	1	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
14	1.1	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
15	1.2	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
16	1.3	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
17	1.4	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187
18	1.5	160.07348	13.22626	30.56404	29.83508	9.97721	10.60187

FIGURE 4 | **An example DVH csv file opened in Excel**. The file is partially displayed.

statistical results were included at the end of each spreadsheet. It took more than 1 h to manually extract and tabulate the same dose/volume points, not including the time of loading the case in DVH viewer of MIMvista.

DISCUSSION

The DPS tool shows great capability of efficiency on dosimetry reviews by automating data extraction. The data extraction

TABLE 2 | Examples of expressions of dose/volume points and their definitions.

Expressions	Definition
D(Gy)40(cc)	The dose in Gy that covers 40 cc of the structure
D(Gy)40(%)	The dose in Gy that covers 40% of the structure
D(%)40(cc)	The dose in percentage that covers 40 cc in the structure
D(%)40(%)	The dose in percentage that covers 40% of the structure
V(cc)40(Gy)	The volume in cc that is covered by a dose not <40 Gy
V(cc)40(%)	The volume in cc that is covered by a dose not <40% of the reference dose
V(%)40(Gy)	The volume in percentage that is covered by a dose not <40 Gy
V(%)40(%)	The volume in percentage that is covered by a dose not <40% of reference dose
Dmax	Maximum dose defined by RTOG, the same as D(Gy)0.03(cc)
Dmin	Minimum dose defined by RTOG, the lowest dose for a point in the volume that is at least 0.03 cc in size (found on the DVH curve at a volume that is the total volume of the structure – 0.03 cc)
Dmean	Mean dose in the structure
Dmean/Rx	Mean dose divided by the prescription dose
Dmean/ Rx × 100	Mean dose divided by the prescription dose and multiplied by 100

request specifications could be exported and imported, thus one can perform the same analysis for future DVH csv files or as needed, which improves task consistency and documentation completeness.

The report from the DPS tool can be used to evaluate the distribution of plan qualities for a particular protocol. Based on this information, it is possible to make the decision to amend a protocol to improve the performance of institutions registering patients. In addition, DPS allows for comparing different ways to evaluate the dosimetry. For example, structure maximum dose can be in the formats of D(Gy)1(%)/Rx, D(Gy)1(%), and D(Gy)0.03(cc). One can add all the three formats of maximum doses to DPS, and the results in the three formats will be tabulated together, which is a helpful feature for protocol development. Furthermore, with minor revision, DPS may perform biomathematical NTCP and TCP model calculations, which will help to understand the biological outcomes of the cases.

CONCLUSION

With the infrastructures and structure name guidelines adopted for clinical-trial data management (24, 25), automated tools, such as DPS, can greatly improve the efficiency and organization of clinical-trial dosimetry review.

AUTHOR CONTRIBUTIONS

All authors contributed to manuscript writing/revisions. YG, JY, HZ, and YX contributed to development of the tool/method. YG, JY, and YX contributed to tool testing. DP and JG made important revisions to the manuscript to achieve publishable quality.

FUNDING

REFERENCES

1. Graham MV, Purdy JA, Emami B, Harms W, Bosch W, Lockett MA, et al. Clinical dose-volume histogram analysis for pneumonitis after 3D treatment for non-small cell lung cancer (NSCLC). *Int J Radiat Oncol Biol Phys* (1999) 45(2):323–9. doi:10.1016/S0360-3016(99)00183-2

2. Emami B, Lyman J, Brown A, Coia L, Goitein M, Munzenrider JE, et al. Tolerance of normal tissue to therapeutic irradiation. *Int J Radiat Oncol Biol Phys* (1991) 21(1):109–22. doi:10.1016/0360-3016(91)90171-Y

3. Grégoire V, Mackie TR. State of the art on dose prescription, reporting and recording in intensity-modulated radiation therapy (ICRU report No. 83). *Cancer Radiother* (2011) 15(6–7):555–9. doi:10.1016/j. canrad.2011.04.003

4. Bekelman JE, Deye JA, Vikram B, Bentzen SM, Bruner D, Curran WJ Jr, et al. Redesigning radiotherapy quality assurance: opportunities to develop an efficient, evidence-based system to support clinical trials – report of the national cancer institute work group on radiotherapy quality assurance. *Int J Radiat Oncol Biol Phys* (2012) 83(3):782–90. doi:10.1016/j.ijrobp.2011.12.080

5. Purdy JA. Quality assurance issues in conducting multi-institutional advanced technology clinical trials. *Int J Radiat Oncol Biol Phys* (2008) 71(1, Suppl):S66–70. doi:10.1016/j.ijrobp.2007.07.2393

6. Palta JR, Deye JA, Ibbott GS, Purdy JA, Urie MM. Credentialing of institutions for IMRT in clinical trials. *Int J Radiat Oncol Biol Phys* (2004) 59(4):1257–9. doi:10.1016/j.ijrobp.2004.03.007

7. Peters LJ, O'Sullivan B, Giralt J, Fitzgerald TJ, Trotti A, Bernier J, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: results from TROG 02.02. *J Clin Oncol* (2010) 28(18):2996–3001. doi:10.1200/JCO.2009.27.4498

8. Goodman KA. Quality assurance for radiotherapy: a priority for clinical trials. *J Natl Cancer Inst* (2013) 105(6):376–7. doi:10.1093/jnci/djt031

9. Dharmarajan KV, Friedman DL, FitzGerald TJ, McCarten KM, Constine LS, Chen L, et al. Radiotherapy quality assurance report from Children's Oncology Group AHOD0031. *Int J Radiat Oncol Biol Phys* (2015) 91(5):1065–71. doi:10.1016/j.ijrobp.2014.11.034

10. FitzGerald TJ, Urie M, Ulin K, Laurie F, Yorty J, Hanusik R, et al. Processes for quality improvements in Radiation Oncology Clinical Trials. *Int J Radiat Oncol Biol Phys* (2008) 71(1, Suppl):S76–9. doi:10.1016/j. ijrobp.2007.07.2387

11. Bekelman JE, Yahalom J. Quality of radiotherapy reporting in randomized controlled trials of Hodgkin's lymphoma and non-Hodgkin's lymphoma: a systematic review. *Int J Radiat Oncol Biol Phys* (2009) 73(2):492–8. doi:10.1016/j. ijrobp.2008.04.058

12. Mendenhall NP, Meyer J, Williams J, Tebbi C, Kessel S, Laurie F, et al. The impact of central quality assurance review prior to radiation therapy on protocol compliance: POG 9426, a trial in pediatric Hodgkin's disease. *Blood* (2005) 106(11):753.

13. FitzGerald TJ, Bishop-Jodoin M, Cicchetti MG, Hanusik R, Kessel S, Laurie F, et al. In reply to Drs. Bekelman and Yahalom: quality of radiotherapy reporting in randomized controlled trials of Hodgkin's lymphoma and non-Hodgkin's lymphoma. *Int J Radiat Oncol Biol Phys* (2010) 77(1):315–6. doi:10.1016/j. ijrobp.2009.12.051

14. Budach V, Zurlo A, Horiot JC. EORTC Radiotherapy Group: achievements and future projects. *Eur J Cancer* (2002) 38(Suppl 4):134–7. doi:10.1016/ S0959-8049(01)00440-3

15. Purdy JA, Perez CA, Poortmans P. *Technical Basis of Radiation Therapy: Practical Clinical Applications.* Berlin; Heidelberg: Springer (2012).

16. Carrie C, Hoffstetter S, Gomez F, Moncho V, Doz F, Alapetite C, et al. Impact of targeting deviations on outcome in medulloblastoma: study of the French Society of Pediatric Oncology (SFOP). *Int J Radiat Oncol Biol Phys* (1999) 45(2):435–9. doi:10.1016/S0360-3016(99)00200-X

17. Abrams RA, Winter KA, Regine WF, Safran H, Hoffman JP, Lustig R, et al. Failure to adhere to protocol specified radiation therapy guidelines was associated with decreased survival in RTOG 9704 – a phase III trial of adjuvant chemotherapy and chemoradiotherapy for patients with resected adenocarcinoma of the pancreas. *Int J Radiat Oncol Biol Phys* (2012) **82**(2):809–16. doi:10.1016/j.ijrobp.2010.11.039

18. Ohri N, Shen X, Dicker AP, Doyle LA, Harrison AS, Showalter TN. Radiotherapy protocol deviations and clinical outcomes: a meta-analysis of Cooperative Group Clinical Trials. *J Natl Cancer Inst* (2013) **105**:387–93. doi:10.1093/jnci/djt001

19. Perez CA, Stanley K, Grundy G, Hanson W, Rubin P, Kramer S, et al. Impact of irradiation technique and tumor extent in tumor control and survival of patients with unresectable non-oat cell carcinoma of the lung. Report by the Radiation Therapy Oncology Group. *Cancer* (1982) **50**(6):1091–9. doi:10.1002/1097-0142(19820915)50:6<1091::AID-CNCR2820500612>3.0.CO;2-0

20. White JE, Chen T, McCracken J, Kennedy P, Seydel HG, Hartman G, et al. The influence of radiation therapy quality control on survival, response and sites of relapse in oat cell carcinoma of the lung. Preliminary report of a Southwest Oncology Group Study. *Cancer* (1982) **50**(6):1084–90. doi:10.1002/1097-0142(19820915)50:6<1084::AID-CNCR2820500611>3.0.CO;2-W

21. Taylor RE, Donachie PH, Weston CL, Robinson KJ, Lucraft H, Saran F, et al. Impact of radiotherapy parameters on outcome for patients with supratentorial primitive neuro-ectodermal tumours entered into the SIOP/UKCCSG PNET 3 study. *Radiother Oncol* (2009) **92**(1):83–8. doi:10.1016/j.radonc.2009.02.017

22. Miralbell R, Fitzgerald TJ, Laurie F, Kessel S, Glicksman A, Friedman HS, et al. Radiotherapy in pediatric medulloblastoma: quality assessment of Pediatric Oncology Group Trial 9031. *Int J Radiat Oncol Biol Phys* (2006) **64**(5):1325–30. doi:10.1016/j.ijrobp.2005.11.002

23. Donaldson SS, Torrey M, Link MP, Glicksman A, Gilula L, Laurie F, et al. A multidisciplinary study investigating radiotherapy in Ewing's sarcoma: end results of POG #8346. *Int J Radiat Oncol Biol Phys* (1998) **42**(1):125–35. doi:10.1016/S0360-3016(98)00191-6

24. Yu J, Straube W, Mayo C, Giaddui T, Bosch W, Ulin K, et al. Radiation therapy digital data submission process for National Clinical Trials Network. *Int J Radiat Oncol Biol Phys* (2014) **90**(2):466–7. doi:10.1016/j.ijrobp.2014.05.2672

25. FitzGerald TJ. A new model for imaging and radiation therapy quality assurance in the National Clinical Trials Network of the National Cancer Institute. *Int J Radiat Oncol Biol Phys* (2014) **88**(2):272–3. doi:10.1016/j.ijrobp.2013.09.030

26. NRG (2016). Available from: https://www.nrgoncology.org/Clinical-Trials

27. DPS (2015). Available from: https://github.com/Hotrank/Dose-volume-Points-Statistics

28. LU001 (2015). Available from: http://www.rtog.org/ClinicalTrials/ProtocolTable/StudyDetails.aspx?study=1327