| **Section and Topic**  | **Item #** | **Checklist item**  | **Location where item is reported**  |
| --- | --- | --- | --- |
| **TITLE**  |  |
| Title  | 1 | PTV margin calculation for Head and Neck patients treated with VMAT: A Systematic Literature Review  |  |
| **ABSTRACT**  |  |
| Abstract  | 2 | Aim: The intent of the review was to identify different methodological approaches used to calculate the PTV margin for head and neck patients treated with Volumetric Arc Therapy (VMAT) and identifying the factors required to calculate the PTV margin. Materials and Methods: A comprehensive, systematic search of related studies was done using MEDLINE, PubMed, CINAHL, ProQuest (Nursing and Allied Health), Scopus and tipsRO. The literature search included studies published between January 2007 till December 2020. Eligibility screening was performed by two reviewers.Results: A total of seven studies were found. All the reviewed studies used the Van Herk Formula to measure the PTV margin. None of the studies incorporated the systematic errors of target volume delineation in the PTV equation. Inter-fraction errors were assessed in all the studies whilst intra-fraction errors were only included in the margin equation in two studies. The studies showed great heterogeneity in the key characteristics, aims and methods. Findings: Since systemic errors from Target Volume (TV) delineation were not considered and not all studies assess intra-fraction errors, PTV margins may be underestimated. Recommendations for studies to determine the effect of TV variance on PTV margins and to compare PTV margins for various formulas.  |  |
| **INTRODUCTION**  |  |
| Rationale  | 3 | Since the PTV margin calculation influences the PTV margin result, the need to perform a systematic literature review were:* To identify what is already known in the literature with regards to PTV margin calculation in the head and neck region.
* To determine the different methods adopted in research studies to calculate this margin for patients treated with VMAT in the

head and neck region.* To identify the factors that need to be considered when calculating the treatment margin.
* To identify ways that reduce the PTV margin size.
 |  |
| Objectives  | 4 | Identify the methods of calculating the PTV margin for head and neck patients treated with VMAT and adapt the most appropriate method locally based on local resources.  |  |
| **METHODS**  |  |
| Eligibility criteria  | 5 |

|  |  |
| --- | --- |
| **Inclusion Criteria** | **Exclusion Criteria** |
| Studies of patients treated with VMAT | Studies of patients not treated with VMAT |
| Image guidance prior delivery of Radiotherapy | No image guidance  |
| Patients treated to the head and neck region only | Patients not treated to the head and neck region |
| PTV margin calculation | Studies that do not calculate PTV margins |
| No age restriction |  |
| Quantitative study | Qualitative studies |
| Prospective or Retrospective |  |
| Observational or Experimental |  |

The review included the following studies: * Studies post 2007
	+ The rationale for excluding studies that were published pre 2007 is that VMAT was first introduced

as a treatment modality in 2007 (Teoh et al., 2011).* Full articles available
* English language studies
 |  |
| Information sources  | 6 | Data was sought by using CINAHL, MEDLINE, PubMed, ProQuest (Nursing and Allied Health), Scopus and tipsRO. Other searches were performed on the institutional library (Hydi) and ScienceDirect platforms. search was done between April, 2020 and December, 2020. |  |
| Search strategy | 7 | An exhaustive search for related research and studies was done through the following combination of keywords:

|  |
| --- |
| * Nasopharyn\*/Nasal cavity
* Oropharyn\*
* Laryn\*/Supraglottis/Subglottis/Glottis
* Hypopharyn\*
* Oral cavity/Mouth/Tongue
* Sinus\*
* Thyroid
* Lymphoma
* Head and Neck
* Set-up/setup/set up
* Error/errors
* VMAT/Volumetric-Modulated Arc Therapy/Volumetric Modulated Arc Therapy/

 RapidArc Therapy* PTV/Planning Target Volume
 |
|
|
|
|
|
|
|
|
|

The asterisks (\*) next to the keywords were used since certain terms can be written in two ways. An example of this would be the keyword Nasophary\* since the asterisk was used to look for nasopharynx and nasopharyngeal search results. Boolean operators were used such as ‘AND’ and ‘OR’ between the keywords and this allowed the combination of words and phrases to retrieve relevant literature from databases. Inverted commas were used on phrases to include all terms. |  |
| Selection process | 8 | A dual independent review of search results. The first phase of the review was to screen for the inclusion and exclusion of studies based on the title and abstract. The second phase of the review was performed by full text reading of the eligible studies selected in the first phase. This process was also performed independently by the two reviewers. Any disagreements by the two reviewers with regards to data suitability was identified and solved through discussions and reaching a consensus agreement. The studies that fit the criteria were included for the systematic review. |  |
| Data collection process  | 9 | Prior to data collection, a pilot test was performed. This approach was taken to ensure that the most useful and relevant information was extracted from the studies, avoiding the need to revisit papers at a later stage. From the pilot test, it was noted that it would be important to add the following parameters to the data extraction sheet:* Specific region of head and neck under investigation
* Imaging protocol
* Type of immobilisation device used
* Radiotherapy prescription

These modifications did not have an impact on the study design.The following quality measurements guided by the PRISMA checklist are the key constructs for structure and organisation purposes for the reviewed papers:* + Title and year of publication
	+ Geographical location where the study was performed
	+ Details of methods (study design, sampling procedure, length of sample follow up, risk of bias)
	+ Sample number (randomly assigned, withdrawal from study or exclusion with reason)
	+ Age range of the sample
	+ Anatomical region of the head and neck
	+ Prescribed dose
	+ Institution PTV margin
	+ Immobilisation
	+ Type of RT linear accelerator and other equipment used
	+ Imaging protocol (frequency, matching procedure, and type of imaging)
	+ Calculated PTV method (statistical analysis)
	+ Reason for choice of calculation method
	+ PTV margin result
 |  |
| Data items  | 10a | The outcome of the study was to determine the various methods for calculating PTV margin for head and neck patients treated with VMAT, and their effect on the PTV margin result. All results that were compatible with each outcome domain in each study were sought. |  |
| 10b | Participants – Studies that analysed Head and neck patients treated with VMATIntervention characteristics – PTV margin calculation |  |
| Study risk of bias assessment | 11 | The Joanna Briggs Institute (JBI) critical appraisal tools were thought to be suitable to assess individual bias in observational studies since these tools can appraise both analytical cross-sectional studies and case control studies.JBI Systematic Reviews Checklist for Analytical Cross-Sectional Studies was concerned with the following factors in the selected studies: clearly defined inclusion and exclusion criteria in a study, clear description of the population of interest,confounding factors, selection bias, reliability and validity of exposed measures and outcome measures, and statistical analysis (Moola et al., 2017a). These sources of bias can threaten the validity of the results of the studies (Viswanathan et al., 2013). Six studies in this review were eligible for this tool.The Joanna Briggs Institute critical appraisal tools for use in JBI Systematic Reviews Checklist for Case Control Studies assessed different criteria than that of the analytical cross-sectional studies. This tool was concerned with the following: comparison of groups, appropriate matching of cases and control, similar criteria for identifying cases and control, reliability and validity of exposed measures and outcomes, similarity in measurement of exposures for cases and control, confounding factors, exposure period and statistical analysis (Moola et al., 2017b). Only one study in this review was found to be eligible for this tool. |  |
| Effect measures  | 12 | The measures include the equation and the methods employed to calculate the CTV-PTV margin for head and neck patients treated with VMAT.To explain the findings of the studies, a narrative synthesis approach was used. |  |
| Synthesis methods | 13a | The intervention characteristics of each study were tabulated and assessed for homogeneity in terms of the PICOS elements.  |  |
| 13b | A Data extraction sheet was prepared to list the key factors of each review. The following quality measurements are the key constructs for structure and organisation purposes for the reviewed papers:* + Title and year of publication
	+ Geographical location where the study was performed
	+ Details of methods (study design, sampling procedure, length of sample follow up, risk of bias)
	+ Sample number (randomly assigned, withdrawal from study or exclusion with reason)
	+ Age range of the sample
	+ Anatomical region of the head and neck
	+ Prescribed dose
	+ Institution PTV margin
	+ Immobilisation
	+ Type of RT linear accelerator and other equipment used
	+ Imaging protocol (frequency, matching procedure, and type of imaging)
	+ Calculated PTV method (statistical analysis)
	+ Reason for choice of calculation method
	+ PTV margin result
 |  |
| 13c | After extracting information with the aid of the data extraction sheet, results were tabulated and charts were used for comparison of results. |  |
| 13d | A narrative synthesis approach was opted since the clinical, methodological and statistical sources were too diverse to be measured with meta-analysis |  |
| 13e | The narrative synthesis explored patterns in the data. It included an investigation of the differences and similarities between the findings of the studies in the review in a systematic way, with a possible logical explanation for the results of the included studies. |  |
| 13f | N/A |  |
| Reporting bias assessment | 14 | The systematic literature review was susceptible to reporting bias since the study was limited to English language studies and this limitation resulted in language bias as other studies which were performed in a non-English language were excluded. |  |
| Certainty assessment | 15 | The JBI tool was used critically appraise the reviewed studies for a quality evaluation assessment.  |  |
| **RESULTS**  |  |
| Study selection  | 16a |  |  |
| 16b | Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded. |  |
| Study characteristics  | 17 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Author,** **Year and****Country** | **Study Design** | **Patients Charact-eristics** | **Head and neck region** | **Imaging protocol** | **Immobilisation device** | **PTV formula** | **PTV margin result** |
| Yin et al., 2013 Southern China | ProspectiveObservationalAnalytical and Cross-sectional | *N* =15Mean age = 44 years (Range = 37-66 years) | Nasopharynx | Daily CBCT | 5-point TP maskHR not specified | VHMF(inter and intra-fraction errors) | ***Total without CBCT correction***ML= 4.1 mmSI = 3.4 mmAP = 3.5 mm ***Total with CBCT correction***ML*=* 1.7 mmSI = 2.2 mmAP = 2.2 MM |
| Oh et al., 2014South Korea | RetrospectiveObservationalAnalytical and Cross-sectional | *N* =35Mean age not specified | not specified | Daily CBCT | 5-point TP mask Individual HR | VHMF(inter-fraction error) | ML = 3.3 mmSI = 2.8 mm AP = 3.7 mm |
| Anjanappa et al., 2017India | RetrospectiveObservationalAnalytical and Cross-sectional | *N* = 20Mean age = 42 years (Range = 14-70 years) | Nasopharynx | Daily 2D KV imaging (KV images taken on alternate days were reviewed) | 4-point TP maskHR not specified | VHMF(inter-fraction error) | ***Clivus level:***ML= 4.0 mm SI=3.2 mmAP=4.4 mm***C3 level:***ML= 5.0 mmSI = 4.4 mmAP = 5.5 mm***C6 level:***ML = 6.9 mmSI = 4.4 mmAP = 6.4 mm |
| Norfadilah et al., 2017 Malaysia | ProspectiveObservationalAnalytical and Cross-sectional | *N* = 8Mean age = 57 years (Range = 23-83 years) | Oral cancer | Daily CBCT | 5-point TP maskMouth BiteHR not specified | VHMF(inter-fraction error) | ***HFW mouthbite*** ML=3.1 mmSI = 2.2 mm AP = 0.8 mm***SYR***ML= 3.8 mmSI = 6.2 mmAP= 5.1 mm |
| Bruijnen et al., 2018Netherlands | ProspectiveObservationalAnalytical and Cross-sectional | *N*=84Mean age =65 years (Range =39-93 years) | NasopharynxOropharynxLarynx | eNAL | 5-point TP mask Individual HR | VHMF(inter and intra-fraction errors) | ***Nasopharynx***S=2.8 mmI = 2.8 mmA = 2.8 mmP = 2.8 mm***Oropharynx***S = 3.0 mmI = 3.1 mmA = 3.0 mmP = 3.0 mm***Larynx***S = 4.0 mmI =3.6 mmA = 3.1 mmP = 3.1 mm***Combined***S = 3.3 mmI = 3.2 mmA = 3.0 mmP = 3.0 mm |
| Deb et al., 2019India | RetrospectiveObservationalAnalytical and Cross-sectional | *N* = 25Mean age not specified | not specified | Daily imaging (eNAL for CBCT & remaining days with 2D PI) | TP mask with shoulder retraction Standard HR | VHMF(inter-fraction error) | ML= 5.6 mmSI = 6.1 mmAP = 4.7 mm |
| Kukolowicz et al., 2020Poland | RetrospectiveObservationalCase-control | *N* = 30Mean age not specified | Nasopharynx and Larynx | Daily EPID | 5-point TP maskStandard HR | VHMF (inter-fractional error) | ***Prior NAL protocol***AP = 4.0 mm,SI = 6.0 mmML = 4.0 mm***NAL protocol***AP = 3.0 mm SI = 2.2 mmML = 3.0 mm |

 |  |
| Risk of bias in studies  | 18 | The following were identified from the JBI Critical appraisal tools:* All selected studies, except for Deb et al. study (2019), specified the inclusion and exclusion criteria in detail.
* All studies except for Deb et al. (2019) study provided sufficient detail on patients’ characteristics.
* Not all studies measured the exposure in a valid and reliable way since in some studies inter- and intra-observer

variability in image matching was not assessed, manual image registration was not performed, and some studies failed to identify how set-up errors were recorded. * Selection of participants was related to both the intervention and outcomes. Participant selection bias was present in some of the

 studies since there were variation in the patient’s characteristics and, at times, lack of information on these characteristics that have a negative effect on the validity of the results.* Outcome measures were not always measured in a valid and reliable way. Some of the studies measured PTV

 margin based on inter-fractional translational errors only and did not consider intra-fractional error, rotation factors, organ motion, and variation in target volume delineation.* Some of the reviewed studies had a small sample size which rendered the results to be unreliable.

|  |  |
| --- | --- |
| **Study** | **Outcome of the evaluation** |
| Oh et al. (2014) | Very strong |
| Bruijnen et al. (2018) | Strong |
| Yin et al. (2013) | Moderate |
| Norfadilah et al. (2017) | Weak |
| Deb et al. (2019) | Weak |
| Anjanappa et al. (2017) | Moderate |
| Kukolowicz et al. (2020) | Strong |

 |  |
| Results of individual studies  | 19 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Study**  | **Target Delineation** | **Intra-fraction error** | **Set-up errors** | **PTV formula** |
| Oh et al. (2014) | x | x |  | PTV= 2.5∑+0.7σ |
| Bruijnen et al. (2018) | x |  |  | PTV = 2.5√(∑motion2+∑setup2) + 0.7√(σmotion2+σsetup2) |
| Yin et al. (2013) | x |  |  | PTV = 2.5√(∑inter-fraction2+∑intra-fraction2) + 0.7√(σinter-fraction2+σintra-fraction2) |
| Norfadilah et al. (2017) | x | x |  | PTV=2.5Σ+0.7σ |
| Deb et al. (2019) | x | x |  | PTV=2.5Σ+0.7σ |
| Anjanappa et al. (2017) | x | x |  | PTV=2.5Σ+0.7σ |
| Kukolowicz et al. (2020) | x | x |  | PTV=2.5Σ+0.7σ |

|  |  |  |
| --- | --- | --- |
| **Study** | **Confounding Variables** | **Strategy** |
| **Oh et al. (2014)** | Intra-fractional movement | Not specified |
|  | Curved external anatomy | Not specified |
|  | Loosening of fixation mask due to weight loss or tightening of mask due to swelling | Thermoplastic mask was remade if considerable discrepancies occurred. Rescanning and replanning were performed when necessary to reduce setup errors. |
| **Bruijnen et al. (2018)** | Accuracy of deformable image registration | Not specified |
|  | Left-right motion affecting image registration | Image acquisition of 10 mm was used. Study referred to previous studies that reported that this motion is small when using this type of acquisition. |
|  | Treatment modality | VMAT PTV margin was calculated by halving the tumour shift between the two cine MR scans |
|  | Persistent tumour motion over a long period of time | Not specified |
| **Yin et al. (2013)** | Weight loss | Examined relationship between weight loss and setup errors and analysed the time trend of weight loss |
|  | Tumour shrinkage | Not specified |
|  | Uncertainty in image registration | Not specified |
|  | Not able to adjust rotational errors | Not specified |
| **Norfadilah et al. (2017)** | No mention | Not specified |
| **Deb et al. (2019)** | Rotation | Not specified |
|  | Weight loss | Not specified |
|  | Tumour shrinkage | Not specified |
| **Anjanappa et al. (2017)** | Rotation | Not specified |
|  | Weight loss | Not specified |
|  | Quality of KV portal imaging and DRR imaging | Not specified |
|  | Difficulty in image registration due to superimposition of bony structures | Not specified |
| **Kukolowicz et al. (2020)** | Rotation | The study mentions that it was very seldom to observe rotations larger than 1 degree, therefore the rotational factor was negligible |
|  | Variation in treatment modality (VMAT and IMRT) | Not specified |
|  | Variation in linear accelerators used | Not specified |
|  | Anatomical changes not visible on portal imaging | Not specified |

 |  |
| Results of syntheses | 20a | The JBI Critical Appraisal tools for use in JBI Systematic Reviews Checklist for Analytical Cross-Sectional Studies was concerned with the following factors in the selected studies: clearly defined inclusion and exclusion criteria in a study, clear description of the population of interest, confounding factors, selection bias, reliability and validity of exposed measures and outcome measures, and statistical analysis |  |
| 20b | N/A |  |
| 20c |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Study** | **Population** | **Intervention** | **Comparative Intervention** | **Outcome** | **Study Design** |
| Oh et al. (2014) | Patients treated with VMAT to the Head and Neck, Brain, Prostate, Thorax and Abdomen | Assessed set-up error and calculated the PTV margin | Compared set-up errors and calculated PTV margin for various tumour sites | To reduce set-up errors and for optimisation of PTV margin | Cross-sectional, retrospective, and quantitative |
| Bruijnen et al. (2018) | Patients treated with IMRT and VMAT to the NasopharynxOropharynxLarynx | Quantified intra-fraction motion and assessed set-up errors | N/A | To determine population based PTV margin | Cross-sectional, prospective, and quantitative |
| Yin et al. (2013) | Patients treated with VMAT to the Nasopharynx | Evaluated inter-fraction and intra-fraction errors | N/A | To determine the set-up errors and appropriate PTV margin | Cross-sectional, prospective, and quantitative |
| Norfadilah et al. (2017) | Oral cancer patients receiving treatment with VMAT | Evaluated and quantify inter-fraction set-up errors for two different immobilisation devices | Compared HeadFIX® mouthpiecemoulded with wax with 10 ml/cc syringe barrel | To determine which immobilisation device produces the least set-up errors | Cross-sectional, prospective, and quantitative |
| Deb et al. (2019) | Head and neck patients treated with VMAT | Assessed set-up error and derived the PTV margin | N/A | To determine the optimal PTV margin | Cross-sectional, retrospective, and quantitative |
| Anjanappa et al. (2017) | Patients treated with VMAT or IMRT to the Nasopharynx | Evaluated inter-fraction set-up errors and derived the PTV margin | Compared the systematic error and random error of three levels: C3, C6 and Clivus | To determine the PTV margin of the Nasopharynx at three different levels | Cross-sectional, retrospective, and quantitative |
| Kukolowicz et al. (2020) | Head and neck patients treated with VMAT or IMRT | Evaluated the impact of NAL imaging protocol with regards to treatment time and set-up errors | Compared the daily imaging protocol with NAL protocol | To reduce treatment time with an effective set-up control | Case-control,both prospective and retrospective, and quantitative |

All the seven studies in the review, except for Kukolowicz et al. (2020), had a quantitative, cross-sectional research design. The population consisted of patients treated to the head and neck region with VMAT. Apart from VMAT, Bruijnen et al. (2018), Anjanappa et al. (2017) and Kukolowicz et al. (2020) mentioned that the patient population also consisted of those that received IMRT.Other variations in the studies laid on the research intervention, comparison, and outcomes. Oh et al. (2014), Deb et al. (2019) and Anjanappa et al. (2017) had similar interventions, that of evaluating and assessing set-up errors, and calculating the PTV margin. Norfadilah et al. (2017) study also had a similar intervention however this study was focused on assessment of set-up errors arising from two different mouthpieces. All these four studies varied in comparison elements and research outcomes.Other similar study interventions were those of Bruijnen et al. (2018) and Yin et al. (2013) since these studies were interested in evaluating and quantifying inter-fraction and intra-fraction errors. These studies also had similar outcomes, that of determining the appropriate PTV margin.Kukolowicz et al. (2020) wanted to assess the effect of the Non-action level (NAL) imaging protocol on the PTV margin and treatment times (Kukolowicz et al., 2020). This study had nothing in common with the other studies in the review with regards to study design since it was a case-control study, intervention, comparative intervention, and outcomes, therefore no comparison could be made with regards to research results.The two studies with the most comparable PICOS elements were those of Bruijnen et al. (2018) and Yin et al. (2013) since they had similar intervention, comparative intervention, outcome, and study design. |  |
| 20d | N/A |  |
| Reporting biases | 21 | The systematic literature review was susceptible to reporting bias since the study was limited to English language studies and this limitation resulted in language bias as other studies which were performed in a non-English language were excluded. Another reporting bias was that of location bias since access to data was limited as the researcher was not able to go through all the resources related to health sciences due to a limitation in time and resources, however performing a dual-independent research design aided in expanding the search |  |
| Certainty of evidence  | 22 | N/A |  |
| **DISCUSSION**  |  |
| Discussion  | 23a | All the studies assessed inter-fraction errors from set-up errors recorded from the imaging software by considering the translational errors. The standard deviation for the population random errors was found to be slightly higher than that of the population systematic errors. The systematic and random errors of set-up rotational errors were considered for some studies however the obtained value was not utilised for the final PTV margin value. Overall, intra-fractional errors were less than inter-fractional errors, and this indicates that immobilisation devices were better at maintaining the position rather than reproducing it.The findings of the review where in line with other studies that stated that different anatomical regions, imaging protocols, immobilisation devices, treatment modality, set-up procedures and patient collaboration have an influence on the size of the PTV margin. |  |
| 23b | The small sample size limits the review in finding significant relationships from the analysed data and this limitation could render the data analysis to be weak.Some of the studies in the review had a weak quality evaluation when evaluated with the Joanna Briggs Institute tool, thereforesome of the studies were not considered to be reliable in terms of outcomes measures and statistical analysis. There was heterogeneity in terms of the key characteristics of the studies and methodology design, therefore this limited comparison of study results.  |  |
| 23c | The systematic review relied on pre-existing data and therefore the results obtained from the data analysis relied on the methodology of the studies in the review. Relying on pre-existing data could introduce a self-reported data bias |  |
| 23d | The clinical implications of the study were to include the evaluation of inter-fraction motion, intra-fraction motion and target volume delineation in the margin calculation using the Van Herk Formula. The departments should also ideally opt for daily imaging as this appears to have a huge impact on the margin size. |  |
| **OTHER INFORMATION** |  |
| Registration and protocol | 24a | registration number: CRD42020183573 |  |
| 24b | The review protocol can be accessed on PROSPERO |  |
| 24c | N/A |  |
| Support | 25 | N/A |  |
| Competing interests | 26 | N/A |  |
| Availability of data, code and other materials | 27 | N/A |  |

*From:*  Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: <http://www.prisma-statement.org/>