Article:
Frood, R, Prestwich, R, Tsoumpas, C orcid.org/0000-0002-4971-2477 et al. (3 more authors) (2018) Effectiveness of Respiratory-gated Positron Emission Tomography/Computed Tomography for Radiotherapy Planning in Patients with Lung Carcinoma – A Systematic Review. Clinical Oncology, 30 (4). pp. 225-232. ISSN 0936-6555

https://doi.org/10.1016/j.clon.2018.01.005

Effectiveness of respiratory-gated PET/CT for radiotherapy planning in patients with lung carcinoma – A Systematic Review

Russell Frood[1], Robin Prestwich[2], Charalampos Tsoumpas[3], Patrick Murray[2], Kevin Franks[2,4], Andrew Scarsbrook[1,4]

1. Department of Nuclear Medicine, Leeds Teaching Hospitals NHS Trust, UK
2. Department of Clinical Oncology, Leeds Teaching Hospitals NHS Trust, UK
3. Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, UK
4. Leeds Institute of Cancer and Pathology, University of Leeds, UK

Corresponding Author:

Dr Russell Frood, Research Fellow, Department of Nuclear Medicine, Level 1, Bexley Wing

St James's University Hospital, Beckett Street, Leeds, LS9 7TF, United Kingdom

Tel: +441132068212

Fax: +441132068228

Email: russellfrood@nhs.net

Running title: 4D PET/CT in lung cancer radiotherapy planning

Word Count: 3499

Number of Figures: 1

Number of Tables: 3

Abstract

Purpose: Systematic review of the literature evaluating clinical use of respiratory-gated (4D) Fluorine-18 fluorodeoxyglucose (FDG) positron emission tomography/computed tomography (PET/CT) compared to non-gated (3D) PET/CT for radiotherapy planning in lung cancer.

Methods: A search of MEDLINE, Cochrane, Web of Science, SCOPUS and clinicaltrials.gov databases was undertaken for articles comparing 3D and 4D PET/CT tumour volume or 4D PET/CT for radiotherapy planning. PRISMA guidelines were followed.

Results: Thirteen studies compared tumour volumes at 3D and 4D PET/CT; 8 reported significantly smaller volumes (6.9% - 44.5%), 3 reported significantly larger volumes at 4D PET/CT (16%-50%), 1 reported no significant difference, and 1 reported mixed findings. Six studies, including 2 which reported differences in tumour volumes, compared target volumes or studied geographic misses. 4D PET/CT target volumes were significantly larger (19%-40%) when compared to 3D PET/CT in all but one study where they were smaller (3.8%). One study reported no significance in 4D PET/CT target volumes when compared to 4D CT, whereas another study reported significantly larger volumes (38.7%).

Conclusion: The use of 4D PET/CT leads to differences in target volume delineation compared with 3D PET/CT. These differences vary depending upon technique and the clinical impact currently remains uncertain. Correlation of pre-treatment target volumes generated at 3D and 4D PET/CT with post-surgical histology would be ideal but technically challenging. Evaluation of patient outcome based on 3D versus 4D PET/CT derived treatment volumes warrant further investigation.

Keywords: Lung cancer; radiation therapy; respiratory gating; FDG PET/CT; systematic review

Introduction

^{18}Fluorine-fluorodeoxyglucose (18F-FDG) positron emission tomography/computed tomography (PET/CT) provides key functional and anatomical information for the staging and management of patients with lung carcinoma, with its role in radiotherapy planning becoming more widely accepted [1–3]. One of the main limitations of thoracic PET/CT is the susceptibility to movement artefact from respiration as, unlike conventional CT imaging, it cannot be acquired in a single breath hold. This can cause blurring of the apparent tumour edge and inaccuracies in measurement of standardised uptake values (SUV) [4]. This may then lead to geographical misalignment of the contoured radiotherapy target volume with the actual tumour position, with the potential for excess normal tissue to be unintentionally irradiated or for geographical misses of the tumour. There is also the theoretical risk that if the patient's breathing patterns are different between follow up scans, the measured change in SUV may be inaccurate and adversely influence the interpretation of treatment response.

Four dimentional (4D) CT is currently the standard-of-care for radiotherapy planning of lung malignancy [5]. Similar methodology has more recently been applied to the use of PET/CT with several methods for gating and contouring of tumours being presented [6]. Studies can be gated by dividing the patient's respiratory cycle and reconstructing the data for either specified amplitude ranges (amplitude-based gating) or specific phase ranges of the respiratory cycle (phase-based gating) [7]. One of the issues currently faced is defining the percentage of the raw data which is included in the recontruction. Too small a percentage of the data will lead to insufficient counts, whereas too great a percentage of the data included predisposes the study to more movement artefact which would nullify the purpose of respiratory gating. This becomes increasingly more difficult when trying to accommodate for irregular breathing patterns [8]. Also, the misalignment of the gated PET and CT data has the potential for inaccuracies in SUV measurement [4][9]. The use of 4D CT for attenuation correction aids in minimising this artefact, however this does increase the radiation dose to the patient

[10]. Another method to aid in registration of respiratory-gated PET and CT is to use a deformation matrix to register all the PET data with respiration, such as a motion freeze technique [11].

The rationale for these methods is to negate respiratory motion, improving accuracy of tumour volume delineation and quantification of lesional tracer activity, potentially enabling more precise metabolillic active tumour targeting [12]. The aim of this article is to systematically appraise the literature and determine whether 4D PET/CT is an effective tool for radiotherapy treatment planning of lung tumours.

Methods

A literature search of MEDLINE/PubMed, Cochrane, Web of Science, Scopus and clinicaltrials.gov databases was performed, searching for articles on the use of 4D PET/CT in lung carcinoma. The search strategy included three major operator criteria which were linked with the "AND" function. The first criteria consisted of "respiratory-gated" or "4D", the second criteria consisted of "PET/CT" or "positron emission tomography" and the third criteria consisted of "lung", "thorax" or "radiotherapy". Case studies, articles not published in English, phantom studies and studies with less than 5 subjects were excluded (to minimise publication bias). After duplications were excluded, studies were screened for eligibility based on title, abstract and subsequently on full text by two authors independently (RF, AS). The results were stored in a bibliographic management software. Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) criteria were adhered to [13]. P values of <0.05 were considered statistically significant.

Results

Results are current to January 2017. The MEDLINE/PubMed, Cochrane, Web of Science, Scopus and clinicaltrials.gov database search strings yielded a total of 1583 results (**Figure 1**). After selection based on review of abstract, the remainding studies underwent full-text assessment. This resulted in 17 articles meeting the inclusion criteria, study characteristics are shown in **Table 1**.

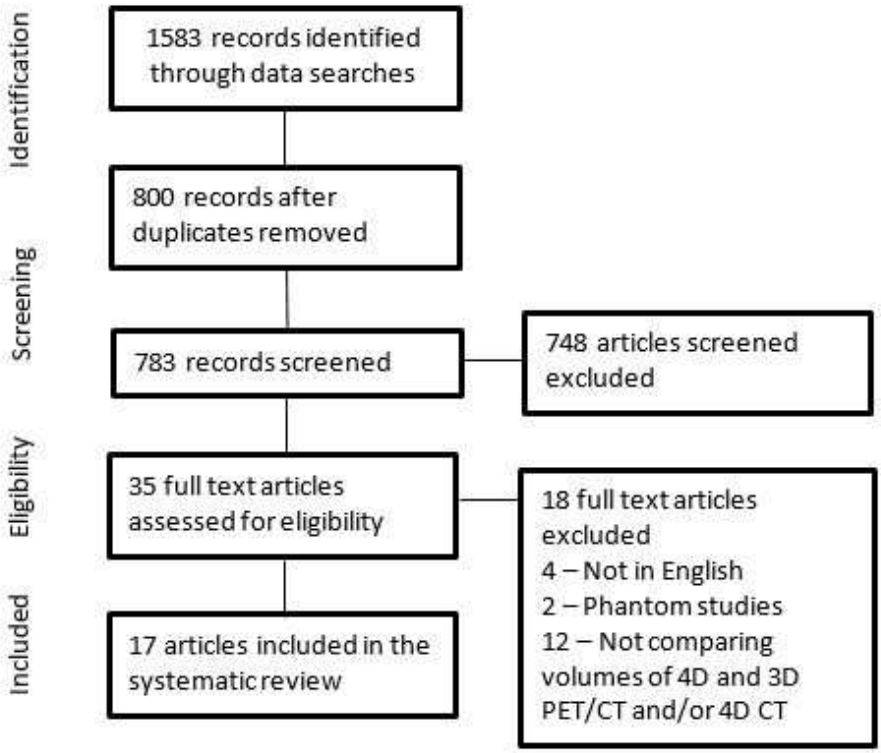

Figure 1: Flow diagram illustrating the methodology for study selection for the systematic review of 4D PET/CT in radiotherapy planning.

Target volumes: 3D PET/CT versus 4D PET/CT

Tradionally when planning radiotherapy, a clinical target volume (CTV) is generated to encompass the gross tumour volume (GTV) and potential areas of adjacent microscopic disease extension. An internal target volume (ITV) can then be created to account for movement of the GTV/CTV within the patient (e.g. due to breathing). In stereotactic ablative radiotherapy (SABR) a CTV is not defined but a composite GTV is drawn on a maximum intensity projection (MIP) then expanded to an ITV directly. A further margin is added to the ITV to account for set-up variability and uncertainties in treatment delivery to create a planning target volume (PTV) [30,31]. By looking at the differences in volumes reported, when using non-gated (3D) PET/CT and 4D PET/CT, it may be possible to determine if firstly, there is a significant difference in the reported tumour volumes between the two methods and secondly if the difference affects the PTV used.

i) Impact on GTV

Thirteen studies out of the 17 included within the literature review assessed the effect of respiratory-gated PET/CT on the GTV when compared to non-gated PET/CT, with mixed results being reported (**Table 2**). Most studies indicate that there is a decrease in measured tumour volume when respiratory-gating is used, but not all of them demonstrated this difference to be significant.

Of the studies which reported a significant decrease in tumour volume, Grootjans *et al.* studied 83 lesions in 66 patients using an optimal respiratory gated (ORG) method [20]. The ORG algorithm determines the amplitude range needed to include a specified proportion of the raw data. The lesions demonstrated a decrease in volume when compared to 3D PET/CT, which was significant in the 20% and 35% duty cycles. When dividing the lesions by location within the thorax there was a significant decrease in volume in the upper lobe and hila lesions in the 20% data cycles and significant decrease in size in the 20% and 35% data cycles in the middle and lower lobe lesions. This suggests that the

amount of tumour motion, which is most prevalent in the lower lobes, accounts for most of the difference between 3D and 4D PET and that using 50% of the data when gating negates the effects. Chang et al. studied a novel way of amplitude gating by performing a CT study under free breathing whilst monitoring the breathing cycle and then only including PET data from that specific amplitude for final image reconstruction [18]. The concept is that the process could be performed on a majority of commercial scanners and the CT and PET should automatically be aligned. They studied 21 lesions and found that the respiratory gated lesions were significantly smaller with average percentage difference being 37.1%. Wijsman et al. also demonstrated significantly smaller GTVs when using amplitude based respiratory gating [29].

Werner et al. studied a phase based method for respiratory gating in 23 lesions [28]. They too found that the GTVs were significantly smaller for the gated PET/CT studies compared to the non-gated study, with the mean tumour volume at 3D PET/CT being 69.0 cm^3 compared to 47.8 cm^3 on 4D PET/CT. Nehmeh et al. also demonstrated smaller GTVs when using a phase based system [24]. Salavati *et al*. compared 3D and 4D PET/CT in 106 lesions, with the gated study being reconstructed in 4 different phases of the respiratory cycle [25]. Unlike the previously described studies they demonstrated no significant difference in GTVs between the non-gated and any of the phases of the respiratory-gated studies. However, as the paper mentions, the data may have been affected by irregularities in patient breathing.

As demonstrated in **Table 2** there are differences in approaches to contouring GTVs between groups. Aritophanous *et al*., who used a phase-gated approach to 4D PET/CT, studied 3 different methods for calculating tumour volume [14]. They evaluated 2 automated contouring protocols, based on >2.5 and GMM (Gaussian Mixture Models), and one manual-based protocol using 40% SUVmax, in 22 lesions

including lymph nodes [14]. The difference in volume for all 3 methods was statistically larger on 4D PET/CT. There was a larger percentage difference in GTVs in lesions which moved more than 3mm (54%) compared to those moving less than 3mm (14%). Callahan *et al*. also demonstrated larger tumour volumes when comparing a 3D and a 4D PET/CT MIP with 4D PET/CT tumour volumes being on average 50% (range 2-446%) larger than the 3D PET/CT ($p < 0.01$) [16].

Van Elmpt *et al.* compared a phased-gated and an optimal-gated amplitude-based system, using 35% of the 4D PET/CT data, to study the use of the different gating methods on the volume and SUV of lesions in 26 lung cancer patients [27]. There were mixed results reported with the phased-based 4D PET showing significantly smaller volumes and higher average SUV when compared to the 3D PET using a 40% threshold of SUVmax to delineate tumour size (P=0.007) but no significant difference when using a 2.5 SUV threshold to determine volume. There was no significant difference in the volume between the optimal-gating and 3D, and no significant difference between the optimal-gated 4D PET and the phased-gated 4D PET. It should be noted that their sample size was smaller (n = 26) than Grootjans et al. (n=83) who used ORG and demonstrated that there was no significant difference in upper and central tumours when using 35% of the data, therefore if a higher percentage of tumours were taken from these areas this may affect the results [20].

The use of a deformation matrix tries to overcome the balance of how much PET data should be included in the gated study to get the best signal to noise ratio. Huang *et al*. looked at 6 patients using 3D and 4D PET/CT applying a motion freeze technique of reconstruction [11]. Five of the lesions were situated within the lower lobes of the lungs and one within the upper lobe. They reported smaller tumour volumes on 4D PET/CT than at 3D PET/CT but with motion freeze demonstrating significantly smaller volumes than both 3D and 4D PET/CT.

Suzawa *et al*. studied the use of respiratory gated time of flight (TOF) PET/CT in the analysis of lung lesions [26]. The principle of TOF PET/CT is that instead of traditional PET/CT where the annihilated positron can be located to a line of response the location can be more accurately located along this line due to calculating the times of arrival of the two 511KeV photons. This in theory allows more accurate imaging requiring fewer counts. Lesions were significantly smaller when compared to the non-gated studies with the percentage difference being greater in lesions which were < 3 cm in size.

There is also variation in the methods used for determining the patient's respiratory cycle which may play a part in the differences in reported volumes. Büther *et al*. explored two different methods of determining the respiratory cycle when gating a cohort of 74 patients with 164 lesions (upper abdominal and thoracic lesions) [15]. Their first method involved using a conventional external pressure sensor whereas the second used information from PET/CT to determine range of movement during the respiratory cycle. There was no significant difference between the volumes of lesions when comparing the two gating methods, but the gated studies produced significantly smaller volumes than the non-gated 100% and 35% data sets. The 2nd (data-driven) method allows gating without additional hardware, is operator independant and produces similar results. Kesner *et al*. also compared hardware and software gating to non-gated PET/CT [23]. They reported no significant difference between volumes calculated using hardware and software gating methods however in their study both methods produced significantly larger tumour volumes when compared to non-gated PET/CT.

One of the main limitations of the majority of these studies is the relatively small numbers of patients leading to inconsistent results and although PET imaging provides excellent image contrast between malignant and normal tissues, the tumour edge is blurred by a combination of limited spatial

resolution and partial volume effects. Phantom studies validate the idea that lesions can be larger or smaller than the true lesion when using respiratory-gating [32]. This may be due to the respiratory cycle of the patient, the size of the lesion (as partial voluming cannot be completely removed by gating) and the amount of data used when constructing the 4D PET/CT, all of which have been shown to under or over-estimate the lesion size [22,32]. In addition, the studies present a variety of methods of tumour segmentation and attenuation correction mapping. There is an extensive literature examining differing methods of segmentation with no clear consensus for the optimal methodology [33]. Therefore, the differing segmentation methods used in Table 1 inevitably impact upon the eventual evaluation of tumour volume. MIP reformatting and >2.5 SUV thresholding are generaly more likely to be associated with larger volumes when compared to other techniques. Also, the distribution and the size of the lesions within the different studies will impact the average volumes. As a small lesion within the periphery of the lung may not demonstrate tracer acitivty due to the confounding partial voluming and motion artefact on the 3D PET/CT but may be demonstrated when the motion artefact is accounted for. This would result in an increase of volume of 100% which could impact on the reported data. A recent systematic review by Sindoni et al. on 4D PET/CT in radiotherapy planning for lung carcinoma discussed the affect of 4D PET/CT on tumour volumes however it did not go into the same depth as our paper or include as many papers in their analysis.

A prospective study evaluating patients who undergo 3D and 4D PET/CT prior to surgical resection has not been performed but would clarify how 4D PET/CT volumes compare to pathological tumour size, similar to the work by Schaefer et al. who compared histological specimens to non-gated PET/CT [34]. This would aid in the validation of 4D PET/CT in radiotherapy planning as there would be more confidence in the volumes provided.

ii) Impact on ITV and PTVs

Although the differences in GTV described above may impact upon radiotherapy planning, it is necessary to consider the ITVs and PTVs; if these target volumes are not significantly different then these reported variations are unlikely to be clinically significant. Respiratory gated CT MIP is currently the technique used to define ITVs and PTVs for radiotherapy planning, and therefore studies examining 4D PET/CT have used these techniques for comparision **(Table 3)**.

Chirindel et al. demonstrated that 4D PET/CT ITVs were significantly larger compared to 4D CT and 3D PET derived ITVs, mean 8.6 cm^3 compared to 6.2 cm^3 for 11 peripherally based lesions [19]. However, all the 4D PET/CT PTVs for the peripheral lesions were incorporated into the PTVs derived from CT. There was no significant difference in the size of central lesions (n=10) between the different techniques, mean volume 44.2 cm^3 compared to 42.1 cm^3 for 4D and 3D PET/CT respectively, however in 2 lesions the PTVs extended further than the PTVs defined on CT and therefore highlights the potential for geographical misses. The difficulty of using CT to define ITV and PTV for central lesions, if IV contrast is not used, is defining the extent of the tumour and therefore the use of respiratory gated PET/CT may have added value in this circumstance. Guerra *et al.* studied differences between 4D PET/CT, 4D CT and 3D CT in 13 patients with solitary lesions [21]. They found no significant difference between 4D PET/CT and 4D CT.

Callahan *et al*. studied geographic misses in radiotherapy planning in 29 patients when comparing 3D and a 4D PET/CT MIP [16]. Four different PTVs were created by adding different margins around the volume: 5 mm, 10 mm, 15 mm and anisotropically 10 mm (laterally) by 15 mm (superiorly-inferiorly). Results were analysed by splitting geographic misses into three types: type 1, any part of the 4D tumour volume outside the 3D target volume; type 2, any part of the 4D target volume outside the 3D

target volume; and type 3, any part of the 4D tumour volume receiving < 95% of the prescribed dose based on the 3D target volume. The number of type 1 misses increased as the PTV margins were decreased with the proportion of type 1 misses also increasing with tumour motion. All PTVs had a type 2 miss. 25/29 cases in the 5, 10, and 15 mm PTV margin groups had a type 3 miss and the asymmetrical margin had one additional miss. There was a significant correlation between lesion motion and percentage of 4D PTV compared to 3D PET/CT with a stronger correlation when the motion to lesion size ratio was used. If there was > 20 mm of movement it was more likely to result in a significant miss (< 90% of the 4D PTV receiving 95% prescribed dose) whereas if there was < 5mm of motion any miss was minor (< 5% of the 4D PTV receiving 95% prescribed dose) and unlikely to be clinically significant. The study concluded that using 3D PET/CT without motion suppression with a PTV margin which is 15 mm or less is proneto more geographic misses especially in lesions with a greater magnitude of motion. A further study by Callahan *et al*. also demonstrated larger ITVs on 4D PET MIP when compared to 3D PET/CT [17].

Wijsman et al. also studied the difference between PTVs derived from 3D and 4D PET/CT [29]. They looked at 22 lesions using a 40% SUV threshold for contouring GTVs for which a 10 mm circumference was added to determine the CTV, organs at risk were subtracted from the CTV, and then a further 5mm circumference was added to determine the PTV. The treatment plan was based on 66 Gy in 33 fractions using a volumetric modulated arc therapy technique with a dose coverage of 99% for the 95% iso-dose coverage. They demonstrated significantly smaller GTVs, CTVs and PTVs at 4D PET/CT when compared to 3D PET/CT. The volume of lung, including the GTV, receiving at least 35 Gy was significantly smaller with 4D PET/CT treatment planning (median difference 5.7cm^3) otherwise there was no significant reduction in the non GTV containing lung or other organs at risk.

Again there is slight variance in the results presented. Jani *et al.* also looked at different methods of gating PET to determine if there was a significant difference in measurement between phase- and amplitude-gated studies when comparing ITVs [22]. They studied 12 lesions and 9 lymph nodes with PET data which was gated into 8 equal amplitude-based bins (A1), equal counts amplitude-based bins (A2) and two temporal-based gating with the windows centred half a phase out from each other. The results showed that amplitude-based methods produced significantly larger ITVs compared to temporal methods with the A1 method producing more accurate volumes in their phantom model correlation.

Recent Developments in Respiratory-Gating Technology

Data-driven or software respiratory gating techniques have recently been developed and are now available on the latest generation PET/CT scanners. These eliminate the need for external hardware to track respiratory motion and involve direct mathematical modelling of the motion of tissues or lesions based on PET acquisition data and have been shown to have equivalent accuracy [15,35]. These recent advances in respiratory gating facilitate automated, operator independent data processing but require standardisation and validation in a multi-centre trial setting before clinical translation. To the best of our knowledge there is no published data on the use of data-driven 4D PET/CT in radiotherapy planning.

Key Recommendations

Key recommodations for further evaluation of 4D PET/CT in radiotherapy planning in lung carcinoma are as follows:

- Focussed study to validate the most appropriate gating methods. The use of software gating should reduce the complexity of data acquisition and allow standardisation of technique ahead of a multi-centre trial(s)
- A prospective study correlating histological tumour volume with tumour volumes delineated on 4D PET/CT and 3D PET/CT would be valuable but may be technically challenging due to a combination of factors including patient preference for non-surgical treatment hampering recruitment and difficulties in accurately comparing metabolic and histological tumour volumes
- A randomised multi-centre trial powered to evaluate clinically relevant efficacy endpoints between a standard 4D-CT radiotherapy planning control arm and 4D PET/CT guided radiotherapy arm using standardised methodology for data acquisition, segmentation and target delineation

These recommendations provide a basis for future translational research in 4D PET/CT guided radiotherapy planning in lung cancer with a greater likelihood of practice changing results.

Conclusion

Due to heterogeneous methodology in the published literature it remains unclear whether 4D PET/CT guides more accurate target volume delineation compared to 3D PET/CT or 4D CT alone in patients with lung cancer. A number of factors have to be considered including spatial resolution of PET, degree of tumour motion, the quantity of data used to reconstruct gated imaging and lesion size. Correlation of pre-treatment target volumes generated at 3D and 4D PET/CT with post-surgical histology would help in design of a future multi-centre trial but may be technically challenging. New data-driven gating methods provide a more realistic technique to implement in evaluation of patient outcomes based on

3D versus 4D PET/CT derived treatment volumes, ideally in the context of prospective randomised study.

References

[1] Kiffer JD, Berlangieri SU, Scott a M, Quong G, Feigen M, Schumer W, et al. The contribution of 18F-fluoro-2-deoxy-glucose positron emission tomographic imaging to radiotherapy planning in lung cancer. Lung Cancer 1998;19:167–77.

[2] Mac Manus MP, Everitt S, Bayne M, Ball D, Plumridge N, Binns D, et al. The use of fused PET/CT images for patient selection and radical radiotherapy target volume definition in patients with non-small cell lung cancer: Results of a prospective study with mature survival data. Radiother Oncol 2013;106:292–8.

[3] Usmanij EA, de Geus-Oei LF, Troost EGC, Peters-bax L, van der Heijden EH, Kaanders JH, et al. PET Early Response Evaluation of Locally Advanced Non – Small Cell Lung Cancer Treated with Concomitant Chemoradiotherapy. J Nucl Med 2013;54:1528–34.

[4] Kruis MF, van de Kamer JB, Vogel W V, Belderbos JS, Sonke JJ, van Herk M. Clinical evaluation of respiration-induced attenuation uncertainties in pulmonary 3D PET/CT. EJNMMI Phys 2015;2:4.

[5] De Ruysscher D, Faivre-Finn C, Nestle U, Hurkmans CW, Le Péchoux C, Price A, et al. European organisation for research and treatment of cancer recommendations for planning and delivery of high-dose, high-precision radiotherapy for lung cancer. J Clin Oncol 2010;28:5301–10.

[6] Sindoni A, Minutoli F, Baldari S, Pergolizzi S. Importance of Respiratory-gated PET/CT in Thoracic Tumors. Radiology 2016;281:321–3.

[7] Tsutsui Y, Kidera D, Taniguchi T, Akamatsu G, Komiya I, Umezu Y, et al. Accuracy of amplitude-based respiratory gating for PET/CT in irregular respirations. Ann Nucl Med 2014;28:770–9.

[8] Son HJ, Jeong YJ, Yoon HJ, Park JH, Kang D-Y. Visual and Quantitative Analysis Methods of

Respiratory Patterns for Respiratory Gated PET/CT. Biomed Res Int 2016;2016:1–11.

[9] Nehmeh SA, Erdi YE, Pan T, Yorke E, Mageras GS, Rosenzweig KE, et al. Quantitation of respiratory motion during 4D-PET/CT acquisition. Med Phys 2004;31:1333–8.

[10] Nehmeh SA, Erdi YE, Pan T, Pevsner A, Rosenzweig KE, Yorke E, et al. Four-dimensional (4D) PET/CT imaging of the thorax. Med Phys 2004;31:3179–86.

[11] Huang TC, Chou KT, Wang YC, Zhang G. Motion Freeze for Respiration Motion Correction in PET/CT: A Preliminary Investigation with Lung Cancer Patient Data. Biomed Res Int 2014;2014:167491.

[12] Konert T, Vogel W, MacManus MP, Nestle U, Belderbos J, Grégoire V, et al. PET/CT imaging for target volume delineation in curative intent radiotherapy of non-small cell lung cancer: IAEA consensus report 2014. Radiother Oncol 2015;116:27–34.

[13] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. BMJ. 2009;339:b2700.

[14] Aristophanous M, Berbeco RI, Killoran JH, Yap JT, Sher DJ, Allen AM, et al. Clinical utility of 4D FDG-PET/CT scans in radiation treatment planning. Int J Radiat Oncol Biol Phys 2012;82:99–105.

[15] Büther F, Vehren T, Schafers KP, Schafers M. Impact of Data-Driven Respiratory Gating in Clinical PET. Radiology 2016;281:229-38.

[16] Callahan J, Kron T, Siva S, Simoens N, Edgar A, Everitt S, et al. Geographic miss of lung tumours due to respiratory motion: a comparison of 3D vs 4D PET/CT defined target volumes. Radiat Oncol 2014;9:291.

[17] Callahan J, Kron T, Schneider-Kolsky M, Dunn L, Thompson M, Siva S, et al. Validation of a 4D-PET maximum intensity projection for delineation of an internal target volume. Int J Radiat Oncol Biol Phys 2013;86:749–54.

[18] Chang G, Chnnag T, Pan T, Clark Jr. J, Mawlawi OR. Implementation of an Automated Respiratory Amplitude Gating Technique for PET/CT: Clinical Evaluation. J Nucl Med 2010;51:16–24.

[19] Chirindel A, Adebahr S, Schuster D, Schimek-Jasch T, Schanne DH, Nemer U, et al. Impact of 4D-18FDG-PET/CT imaging on target volume delineation in SBRT patients with central versus peripheral lung tumors. Multi-reader comparative study. Radiother Oncol 2015;115:335–41.

[20] Grootjans W, de Geus-Oei LF, Meeuwis AP, van der Vos CS, Gotthardt M, Oyen WJ, et al. Amplitude-based optimal respiratory gating in positron emission tomography in patients with primary lung cancer. Eur Radiol 2014;24:3242–50.

[21] Guerra L, Meregalli S, Zorz A, Niespolo R, De Ponti E, Elisei F, et al. Comparative evaluation of CT-based and respiratory-gated PET/CT-based planning target volume (PTV) in the definition of radiation treatment planning in lung cancer: Preliminary results. Eur J Nucl Med Mol Imaging 2014;41:702–10.

[22] Jani SS, Robinson CG, Dahlbom M, White BM, Thomas DH, Gaudio S, et al. A comparison of amplitude-based and phase-based positron emission tomography gating algorithms for segmentation of internal target volumes of tumors subject to respiratory motion. Int J Radiat Oncol Biol Phys 2013;87:562–9.

[23] Kesner AL, Lind KE, Kwak JJ, Lynch D, Burckhardt D, Koo PJ, et al. Validation of Software Gating : A Practical Technology for Respiratory Motion Correction in PET. Radiology 2016;281:239-48.

[24] Nehmeh SA, Erdi YE, Ling CC, Rosenzweig KE, Schoder H, Larson SM, et al. Effect of respiratory gating on quantifying PET images of lung cancer. J Nuc Med 2002;43:876–81.

[25] Salavati A, Borofsky S, Boon-Keng TK, Houshmand S, Khiewvan B, Saboury B, et al. Application of Partial Volume Effect Correction and 4D PET in the Quantification of FDG Avid Lung Lesions. Mol Imaging Biol 2014;17:140–8.

[26] Suzawa N, Ichikawa Y, Ishida M, Tomita Y, Nakayama R, Sakuma H. Respiratory-gated time-of-flight PET/CT during whole-body scan for lung lesions: feasibility in a routine clinical setting and quantitative analysis. Ann Nucl Med 2016;30:722–30.

[27] Van Elmpt W, Hamill J, Jones J, De Ruysscher D, Lambin P, Öllers M. Optimal gating compared to 3D and 4D PET reconstruction for characterization of lung tumours. Eur J Nucl Med Mol Imaging 2011;38:843–55.

[28] Werner MK, Parker JA, Kolodny GM, English JR, Palmer MR. Respiratory gating enhances imaging of pulmonary nodules and measurement of tracer uptake in FDG PET/CT. Am J Roentgenol 2009;193:1640–5.

[29] Wijsman R, Grootjans W, Troost EG, van der Heijden EH, Visser EP, de Geus-Oei LF, et al. Evaluating the use of optimally respiratory gated 18F-FDG-PET in target volume delineation and its influence on radiation doses to the organs at risk in non-small-cell lung cancer patients. Nucl Med Commun 2016;37:66–73.

[30] International Commission of Radiation Units and Measurements. Prescribing, recording, and reporting photon beam therapy (Supplement to ICRU Report 50). ICRU Report 62. Oxford University Press, Oxford, UK (1999)

[31] International Commission of Radiation Units and Measurements. Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU Report 50). ICRU Report 62. Oxford

University Press, Oxford, UK (1999)

[32] Park SJ, Ionascu D, Killoran J, Mamede M, Gerbaudo VH, Chin L, et al. Evaluation of the combined effects of target size, respiratory motion and background activity on 3D and 4D PET/CT images. Phys Med Biol 2008;53:3661–79.

[33] Bhatnagar P, Subesinghe M, Patel C, Prestwich R, Scarsbrook AF. Functional imaging for radiation treatment planning, response assessment, and adaptive therapy in head and neck cancer. Radiographics 2013;33:1909–29.

[34] Schaefer A, Kim YJ, Kremp S, Mai S, Fleckenstein J, Bohnenberger H, et al. PET-based delineation of tumour volumes in lung cancer: Comparison with pathological findings. Eur J Nucl Med Mol Imaging 2013;40:1233–44.

[35] Kesner AL, Schleyer PJ, Büther F, Walter M a, Schäfers KP, Koo PJ. On transcending the impasse of respiratory motion correction applications in routine clinical imaging – a consideration of a fully automated data driven motion control framework. EJNMMI Phys 2014;1:8.

Authors	Year	Number of lesions	Brief Summary
Aristophanous et al. [14]	2012	22*	4D PET/CT GTVs were significantly larger when compared to 3D PET/CT volumes using different thresholds for contouring
Büther *et al*. [15]	2016	164**	4D PET/CT GTVs were significantly smaller compared to 3D PET/CT volumes when using data-driven or belt driven gating
Callahan et al. [16]	2014	29	Using the standard 3D PET/CT 15mm PTV without motion suppression results in more geographic misses
Callahan et al. [17]	2013	9	4D PET MIP GTVs were significantly larger when compared to 3D PET
Chang et al. [18]	2010	21	4D PET/CT GTVs were significantly smaller when compared to 3D PET/CT
Chirindel et al. [19]	2015	21	4D PET/CT ITVs were significantly larger when compared to 4D CT
Grootjans et al. [20]	2014	83	Optimal gated 4D PET/CT GTVs were significantly smaller when compared to 3D PET/CT however, significance depended on the location of lesion and the amount of data used for reconstruction
Guerra et al. [21]	2014	13	4D PET/CT PTVs were larger but this did not reach significance
Huang et al. [11]	2014	6	4D PET/CT GTVs were significantly smaller when compared to 3D PET/CT when using data driven gating
Jani et al. [22]	2013	21*	ITVs using amplitude based 4D gating were significantly larger than those delineated when using phase based gating
Kesner et al. [23]	2016	116***	Both data driven and hardware driven 4D gated PET/CT resulted in significantly larger GTVs when compared to 3D PET/CT
Nehmeh et al.[24]	2002	5	Smaller GTVs with 4D PET/CT when compared to 3D PET/CT
Salavati *et al*. [25]	2014	106	No significance between 4D PET/CT and 3D PET/CT volumes
Suzawa et al. [26]	2016	50	4D PET/CT GTVs were significantly smaller than 3D PET/CT volumes
Van Elmpt et al.[27]	2011	26	4D PET/CT GTVs were only significantly smaller than 3D PET/CT when using phase based gating and 40% threshold contouring
Werner et al. [28]	2009	23	Optimal gating 4D PET/CT produced significantly smaller GTVs when compared to 3D PET/CT
Wijsman et al. [29]	2016	22	4D PET/CT GTVs were significantly smaller when compared to 3D PET/CT volumes

Table 1: Summary of studies included in the systematic review. Key: GTV = Gross tumour volume, ITV = Internal Target Volume, PTV = Planning Target Volume. *Includes lymph nodes, **includes abdominal lesions, ***Number of patients.

Authors	Number of lesions	Type of gating	Contouring	Average volume of lesion on 3D PET (cm^3)	Difference in average tumour volume (GTV) (4D PET/CT vs 3D PET/CT)	Significance	P-value
Aristophanous et al. [14]	9*	Phase	>2.5 SUV	26.1	21.8% larger	Yes	<0.05
			GMM	19.8	19% larger	Yes	<0.05
			MAN 40% SUV max	9	13.5% larger	Yes	<0.05
Callahan et al. [16]	29	Phase (MIP)	MAN	15.7	50% larger	Yes	<0.01
Chang et al. [18]	21	Amplitude	40% SUVmax	11	37.1% smaller	Yes	<0.05
Grootjans et al. [20]	83	Amplitude 20%	40% SUVmax	18.7	11.3% smaller	Yes	<0.0001
		35%		18.7	8.5% smaller	Yes	<0.0001
		50%		18.7	6.9% smaller	Yes	0.02
Huang et al. [11]	6	MF	42% SUVmax	75.7	18.5% smaller	Yes	
		Phase		80.5	13.4% smaller	Yes	
Suzawa et al. [26]	50	Phase	PET Edge	14.7	14.2% smaller	Yes	<0.001
Van Elmpt et al.[27]	26	Phase	>2.5 SUV	74.2	1.2% larger	No	
		Phase	40% SUVmax	32.2	5.3% smaller	Yes	<0.05
		Amplitude	>2.5 SUV	74.2	0.9% larger	No	
		Amplitude	40% SUV max	32.2	2.5% smaller	No	
Werner et al. [28]	23	Phase	42% SUV max	69	44.5% smaller	Yes	<0.05
Wijsman et al. [29]	22	Amplitude	40% SUV max	5.8 (median)	20.7% smaller	Yes	<0.05

Table 2: Percentage difference of average size of lesions evaluated in 3D PET/CT compared to 4D PET/CT reported by different studies. Automatic contouring was performed unless MAN is stated. Key: MAN = manual contouring, PET EDGE = software contouring using steepest change in intensity to determine contour, ACOT = adaptive contrast- oriented thresholding algorithm, MIP = maximum intensity projection, MF = motion freeze, NS = non-significant *Lymph nodes not included in result. Four studies not included due to missing methodology or data required to calculate percentage difference.

Authors	Number of lesions	ITV/PTV	Comparison with 3D PET/CT or 4D CT	Percentage difference	Significance	P-value
Callahan et al. [16]	29	PTV 5mm	3D PET/CT	40% larger	Yes	0.0013
		PTV 10mm	3D PET/CT	32% larger	Yes	0.0001
		PTV 15mm	3D PET/CT	31% larger	Yes	<0.0001
		Asym	3D PET/CT	31% larger	Yes	<0.0001
		PTV 15x10mm	3D PET/CT	19% larger	Yes	<0.0001
Callahan et al. [17]	9	ITV	3D PET/CT	40% larger	Yes	0.0006
Chirindel et al. [19]	21	ITV	4D CT	38.7% larger	Yes	<0.05
Guerra et al. [21]	13	PTV	4D CT	3.4% larger	No	0.16
Wijsman et al. [29]	22	PTV	3D PET/CT	3.8% smaller	Yes	0.036

Table 3: Percentage difference of average size of PTV/ITV evaluated in 3D PET/CT or 4D CT compared to 4D PET/CT reported by different studies.