Validation of RAH VQ SPECT/CT lobar quantification program using a modified version of GE Q lung

Chong Ghee Chew[1] MBBS, FAANMS, FRACP,
Benjamin Crouch[1] BSc, Msc,
Stirling Ha[1] Bmed RadSc (NucMed),
Alice Burke[1] BMed RadSc (NucMed),
Jennie Louise[2] PhD, M.Biostat

1. Department of Nuclear Medicine, PET and Bone Mineral Densitometry, Royal Adelaide Hospital
2. Faculty of Health Sciences, University of Adelaide

Keywords: VQ SPECT/CT
- Lobar Quantification
- Emphysema

Corresponding author:
Chong Ghee Chew MBBS, FAANMS, FRACP
Department of Nuclear Medicine, PET and Bone Mineral Densitometry,
Royal Adelaide Hospital
Ghee.Chew@sa.gov.au

Received:
16 July 2022
Accepted revised:
13 January 2023

Abstract

Objective: The value of ventilation-perfusion (VQ) single photon emission tomography/computed tomography (SPECT/CT) lobar quantification for pre-operative assessment of lobectomy and lung volume reduction is known. Our in-house developed software, RAH ventilation perfusion SPECT/CT quantification (RAHVQSQ) has been shown to be able to identify the target lobe for collapse in bronchoscopic lung volume reduction (BLVR) for advanced emphysema. We have proven inter and intra observer reproducibility but are yet to validate the accuracy of our program. This study aims to validate the accuracy of our quantitative program through comparison with a modified version of GE Q lung which is a commercial program certified for clinical use. **Subjects and Methods:** Ventilation-perfusion SPECT/CT data of 19 subjects from our previous study using RAHVQSQ for BLVR assessment were re-analysed using Q lung by 2 technologists independently and in a blinded fashion to determine lobar differential ventilation, perfusion and volume percentages. The data were from GE Hawkeye 4 and external CT, thus a modified version of Q lung was used. To determine interobserver variation in the 3 parameters between the 3 assessors, intraclass correlation coefficient (ICC) and Bland-Altman limits of agreement (LoA) were generated. **Results:** Paired comparisons between the 3 assessors had high ICC (range for ventilation: 0.69-0.97; perfusion: 0.69-0.97; volume: 0.63-0.97) and means of LoA differences close to zero (range for ventilation: -0.04 - 0.10; perfusion: 0.00-0.02; volume: -0.12 - 0.09) were noted indicative of good concordance for all parameters. **Conclusion:** Using VQ SPECT/CT data of participants with advanced airway disease, our study has found a close concordance of estimated differential lobar ventilation, perfusion and volume percentages using RAHVQSQ when compared with a duplicated blinded assessment using Q lung. The good concordance supports the validity of our quantitative methodology.

Hell J Nucl Med 2023; 26(1): 9-19 *Epub ahead of print: 11 April 2023* *Published online: 28 April 2023*

Introduction

Ventilation-perfusion planar quantification has been in use for assessment of patients being considered for lobectomy and lung volume reduction [1, 5]. This technique involves quantifying the radioactivity of both lungs acquired anteriorly and posteriorly, segmenting the lungs into upper, middle and lower zones, and applying geometric mean to determine the percentage activity in each rectangular segment. The rectangular segments however do not match lobar anatomy and at best this technique provides a very rough estimate of lobar differential function.

Ventilation-perfusion single photon emission tomography/computed tomography (VQ SPECT/CT) has been shown to significantly improve accuracy for assessment of pulmonary embolism [6]. One of the strengths of tomographic and hybrid imaging is improved qualitative delineation of lobar anatomy. This has translated to the added ability for lobar quantification of ventilation and perfusion from the VQ scan and the lobar volume from the CT. There are now several commercial software platforms that can perform lobar quantification. One widely used example is Q lung (GE Healthcare). Ventilation-perfusion SPECT/CT lobar quantification for pre operative assessment for lung volume reduction surgery has been shown to be of value [3, 4].

In our institution we have devised our own lobar quantification program which is written in IDL (Interactive Data Language). We have termed the program Royal Adelaide Hospital VQ SPECT/CT quantification (RAHVQSQ). In addition we have taken a novel approach to the data interpretation by devising quantitative indices that incorporate the lobar ventilation, perfusion and volume with the aim of estimating the differential lobar contribution to total lung function, the differential lobar lung parenchymal injury and the differential lobar gas exchange efficiency [2, 7]. In a clinical study of patients with end stage emphysema being considered for lung volume reduction with endobronchial valves we have shown that lobe selection for collapse to improve mechanical ventilation using one of our indices was highly concordant with the gold standard imaging test of quantitative CT [2]. However, to validate the accuracy of our quantitative technique we are hampered by the lack of an independent gold standard. As a compromise a currently certified quantification method such as GE Q lung will have to serve as a surrogate gold standard for the purpose of validating our methodology.

Aim

The aim of our study is to validate the accuracy of our lobar quantification technique by comparing our results with Q lung as the latter has been certified as a clinical tool. We chose to use technically challenging data from our lung volume reduction study which comprised participants with advanced emphysema. For a more robust comparison, the study data were independently and blindly analysed using Q lung by 2 nuclear medicine technologists.

Subjects and Methods

From our previous clinical study using VQ SPECT lobar quantification with RAHVQSQ to determine the target lobe for collapse using endobronchial valves deployed via bronchoscopy, we had data from 19 participants with end stage emphysema and who were symptomatic despite maximum medical therapy. The baseline demographics and clinical characteristics of the participants are fully outlined in our previous publication [2].

As per the initial study, the VQ SPECT/CT datasets were anonymised and identified with a code. The VQ were performed on GE Hawkeye 4 SPECT/CT (GE Healthcare). Standard protocol was applied with 40-50MBq technetium-99m (^{99m}Tc) Technegas for the ventilation scan and 200-250MBq ^{99m}Tc-labelled macro aggregated albumin for the perfusion scan. Single photon emission tomography scans were acquired for both studies (Matrix 128×128; rotation: 3.0, zoom 1; ventilation 13s/step, perfusion 8s/step; 120 views; Butterworth filter, threshold 0.48, power 10 and OSEM 2 iterations 10 subsets). If a low-dose CT was used, the acquisition parameters were 2.5mA, 140kV, DLP 125mGy cm.

The participants had their imaging in various centres depending on the point of recruitment. Most of the CT were performed on a 128 slice Somatom Definition AS (Siemens Healthcare) with acquisition parameters of 120kVp, 85mAs, DLP 200-300mGy cm, CTDI 5-10mGy, acquired both in inspiration and expiration.

The SPECT and the diagnostic CT datasets were then used for lobar segmentation in the RAHVQSQ program, which is IDL-based version 6.3 software. In this program, the CT is displayed as a 3-dimensional map which can be freely rotated by the assessor who determines the positions of the fissures by manual annotation (Figure 1). The program allows the free manipulation of the windowing of the CT to assist fissure identification. Once the fissures are localized, the program then fuses the CT map with the V and Q SPECT and determines for each lobe the percentage radioactivity of ventilation and perfusion from the VQ scan and the percentage volume from the CT. All of the studies were analysed with RAHVQSQ by one nuclear physician (principal investigator, CC).

Q lung is a program that quantifies percentage lobar ventilation, perfusion and volume of VQ SPECT/CT acquired with a GE hybrid scanner (Discovery 670 or later models). GE Healthcare had kindly provided us a version of Q lung capable of performing the same quantification on VQ data acquired from a GE Hawkeye 4, paired with a CT acquired separately. The VQ data were analysed by 2 nuclear medicine technologists (our co investigators, SH and AB), blinded to the participants' identity, to each other's results and to the RAHVQSQ result. During the process the technologists recorded any technical difficulties experienced.

Statistical analysis

All statistical analyses were performed using Stata v16 (StataCorp, College Station, TX). Because one assessor evaluated patients using RAHVQSQ, while two other assessors evaluated patients using Q lung, it was not possible to separate assessor effects from the effect of the different softwares. Therefore, RAHVQSQ was compared separately with each of the Q Lung assessments, and the two Q Lung assessments were also compared with each other to assess the degree of inter-rater reliability for this method.

Initial analyses were performed to calculate the consistency in the measures of perfusion, ventilation and volume between the methods. The 'consistency-of agreement' intra-

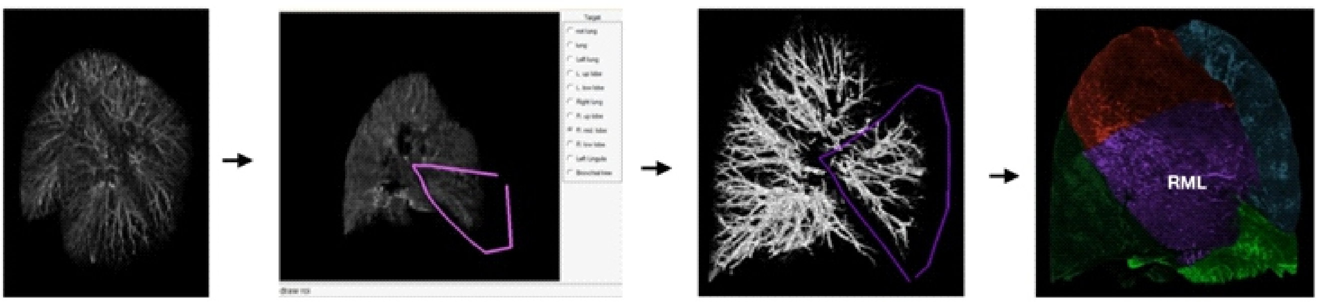


Figure 1. The assessor rotates the image to the desired angle, specifies the target region by annotating a closed curve around the target region which is projected through the whole image. The CT window is adjustable to display the pulmonary vasculature thus accentuating the fissures.

class correlation coefficient (ICC) and 95% confidence interval from a mixed-effects model was estimated, where method (RAHVQSQ; Q lung 1; Q lung 2) was a fixed effect and lobe of the lung was a random effect. However, additional analyses were then carried out to estimate the Bland-Altman 95% limits of agreement, to provide a fuller picture of the extent of agreement between the two methods. The 95% interval for the limits of agreement estimates the interval between which 95% of differences between future measurement pairs would be expected to fall.

Because the standard Bland-Altman method does not account for the clustering in the data (potential correlation between the 5 lobes measured on one patient), a sensitivity analysis was also carried out using the method described in Carstensen et al. (2008) to calculate 95% limits of agreement accounting for the clustering [9].

Results

Analyses of paired concordance between the 3 assessors are depicted in Figure 2. For ventilation, perfusion and volume percentages, the ICC indicates a strong concordance between RAH vs Q1 and RAH vs Q2. ICC for Q1 vs Q2 indicate a very strong concordance which imply Q lung to be a highly reproducible technique. For all parameters, the Bland Altmann limits of agreement estimates show mean differences between the paired comparisons close to zero implying good concordance. For Q1 vs Q2 the limits of agreement were narrower than comparisons with RAH. Sensitivity analyses accounting for clustering produced 95% limits of agreement that were similar to those obtained using the standard Bland-Altman method (data not shown).

When processing the studies using Q lung, the 2 technologists documented any technical difficulties encountered. Table 1 outlines the problems encountered in 10 data points. The nature of the problems was similar, relating to deficient fissures and the program localization of fissure being incongruent with their visual assessment. Of note, among these 10 participants only 2 had discordant results between RAH vs Q1 and RAH vs Q2. With regards to the data of one of these 2 discordant results, for participant 8 both assessors Q1 and Q2 thought the left oblique fissure picked by Q lung was incorrect based on the qualitative assessment. For the data of the other discordant dataset of participant 6, qualitatively assessor Q1 disagreed with the location of the left oblique fissure determined by Q lung. In all 10 participants, Q1 was concordant with Q2.

A second analysis of the paired concordance of the data was performed with the ten data points excluded (Figure 3). The paired comparison between RAH and Q1 showed improved ICC to very strong concordance for ventilation, perfusion and volume percentages. The ICC scores for the other 2 assessor pairs did not change. The limits of agreement did not change significantly.

Discussion

We have shown in a separate blinded study that VQ SPECT/CT lobar quantification using RAHVQSQ can identify a suitable lobe for reduction using endobronchial valves in a cohort of end stage emphysema participants with statistically significant measured clinical improvement [2]. In this study we showed that our choice of target lobe for collapse using RAHVQSQ was 89% (Kappa=0.85) concordant with the use of quantitative CT which is the current gold standard imaging for preoperative assessment. We have used the data from this previous study for the current validation study.

As the purpose of the current study was to validate our quantification program by a direct comparison with a commercial program we have not included details of participant demographics, methodology and measured clinical outcomes of our previous study as these parameters are accessible from our past publication. Moreover they are not pertinent to the current aim. One reason for choosing to use the data sets from this study was because the analyses with RAHVQSQ were already done in a blinded fashion.

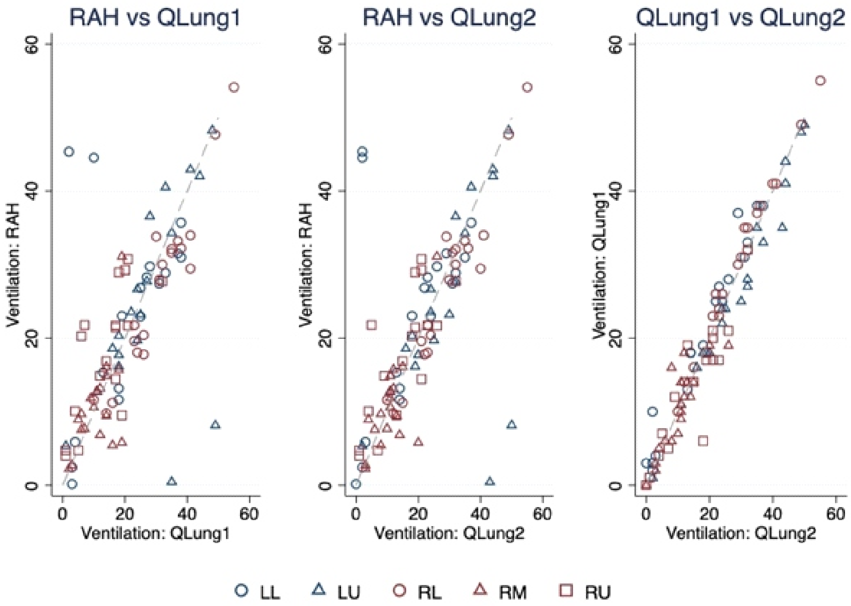
We have also shown in a pilot study that VQ SPECT/CT lobar quantification with our technique can potentially be used to evaluate lobar lung reserve in participants being considered for lobectomy [7]. In this preliminary study the measured 6 months post lobectomy FEV1 and DLCO showed good concordance with the preoperative predicted results extrapolated from the expected residual lung reserve using indices derived from our methodology of differential lobar contribution to total lung function.

We have shown that our methodology is reproducible. In a blinded study using VQ data from participants with advanced airways disease being considered for lung transplantation analysed with RAHVQSQ by 3 assessors we found that our method has high inter and intra observer concordance [8].

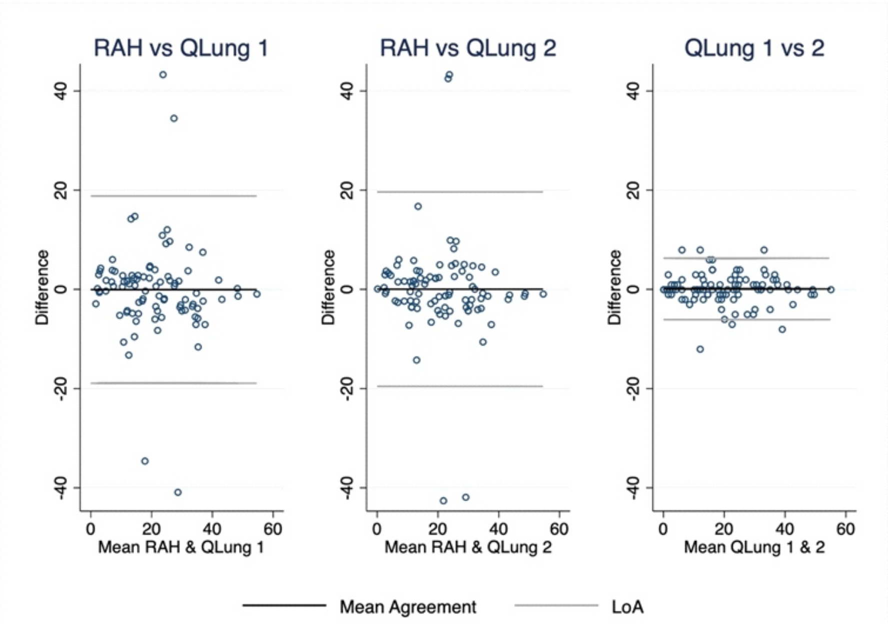
Having shown clinical utility and reproducibility of our method, our aim with this study is to validate our quantification method against a commercially available program that has been approved as a clinical device. We chose to do so because to our knowledge there is no gold standard test for scintigraphic lobar quantification. The data we used are from the participants with end stage emphysema and were being considered for bronchoscopic lung volume reduction for relief of residual symptoms despite maximum medical treatment. The advanced level of parenchymal damage and duplicated blinded assessment with Q lung serve to increase the robustness of this validation study. We found that the RAHVQSQ quantified ventilation, perfusion and volume percentages of the 5 lobes had good concordance with the commercial program.

From a technical perspective we noted a few benefits of our program over Q lung. Our technique utilises a 3D reconstructed CT image that is window adjustable. This was found to be helpful in the event of markedly scarred or collapsed lobes and when the fissures were markedly deficient or near absent.

We also found Q lung to be less flexible in terms of defining the location of the fissures when there is a clear discrepancy with the qualitative assessment. In one of the datasets that there was discordance with RAHVQSQ, both assessors Q1 and Q2 disagreed with the positioning of the left oblique fissure determined by Q lung. In the other discordant dataset assessor Q1 disagreed with Q lung with regards to the left oblique fissure location. Potentially, if both assessors were able to override Q lung for these 2 datasets, there may have been higher

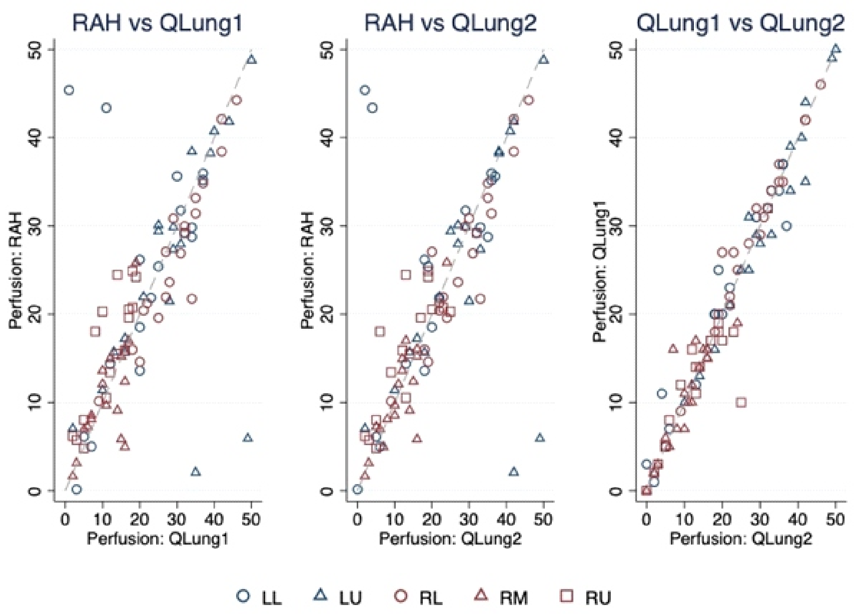


Ventilation		
Comparison	Estimated ICC (95% CI)	Concordance
RAH vs Q1	0.71(0.59, 0.80)	Strong
RAH vs Q2	0.69 (0.57, 0.79)	Strong
Q1 vs Q2	0.97 (0.96, 0.98)	Very Strong

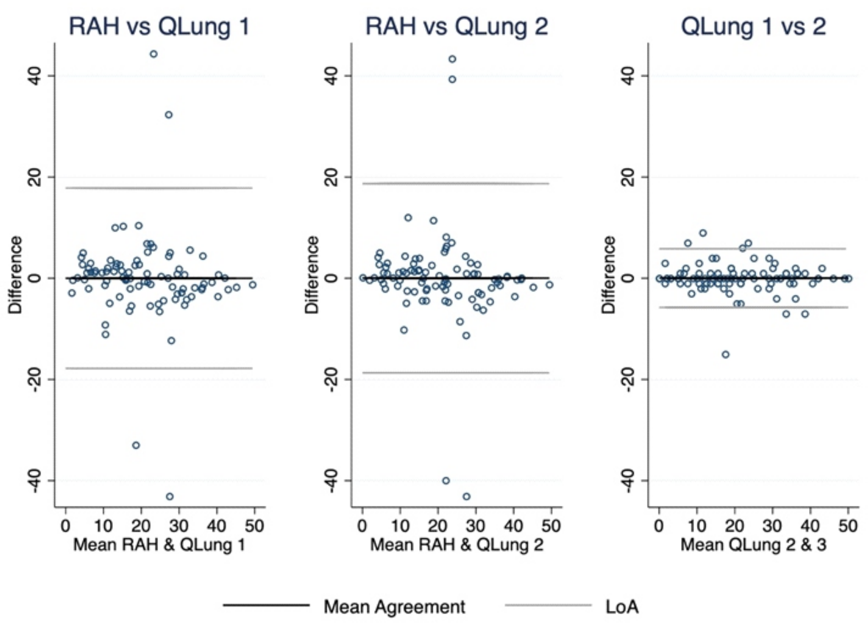


Ventilation	
Comparison	Estimated Mean Difference + Limits of Agreement (LoA)
RAH vs Q1	-0.04 (-18.90, 18.81)
RAH vs Q2	0.05 (-19.55, 19.66)
Q1 vs Q2	0.10 (-6.07, 6.27)

Figure 2a. Lobar differential ventilation concordance by ICC and by mean difference/limits of agreement.

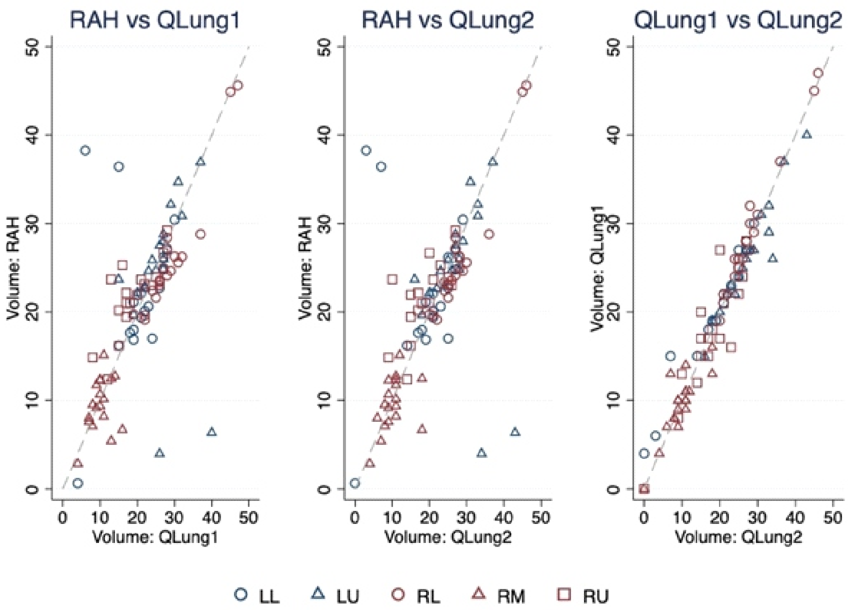


Perfusion		
Comparison	Estimated ICC (95% CI)	Concordance
RAH vs Q1	0.72 (0.60, 0.80)	Strong
RAH vs Q2	0.69 (0.57, 0.79)	Strong
Q1 vs Q2	0.97 (0.96, 0.98)	Very Strong

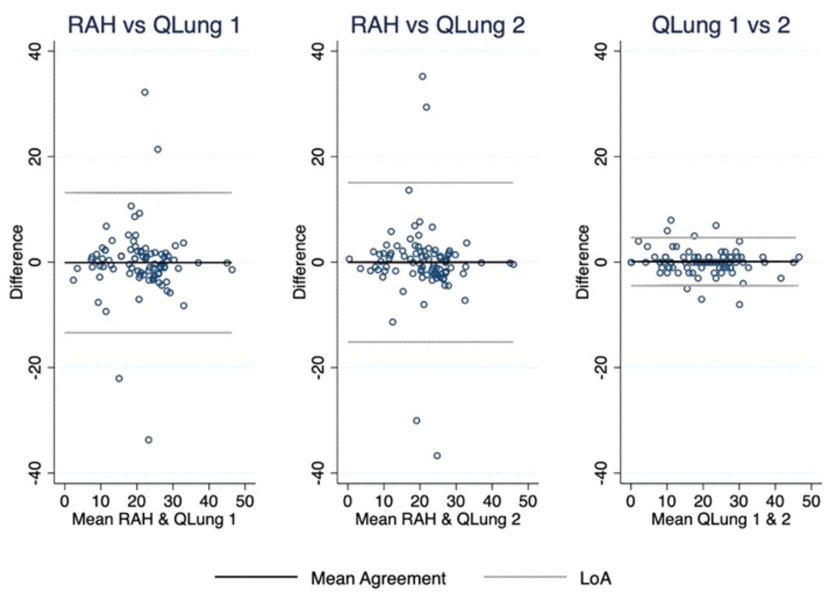


Perfusion	
Comparison	Estimated Mean Difference + Limits of Agreement (LoA)
RAH vs Q1	-0.00 (-17.81, 17.81)
RAH vs Q2	0.02 (-18.68, 18.72)
Q1 vs Q2	0.02 (-5.77, 5.81)

Figure 2b. Lobar differential perfusion concordance by ICC and by mean difference/limits of agreement.



Volume		
Comparison	Estimated ICC (95% CI)	Concordance
RAH vs Q1	0.70 (0.58, 0.79)	Strong
RAH vs Q2	0.63 (0.49, 0.74)	Strong
Q1 vs Q2	0.97 (0.95, 0.98)	Very Strong

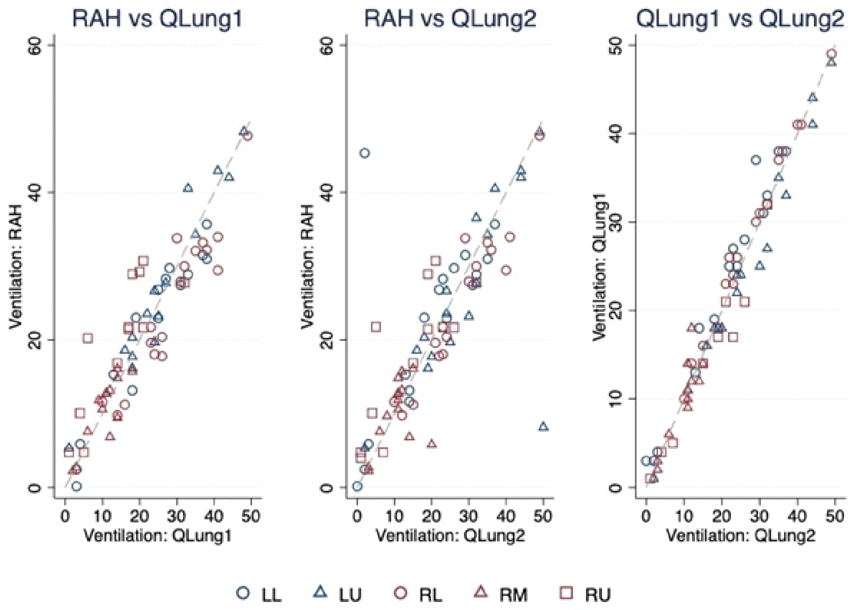


Volume	
Comparison	Estimated Mean Difference + Limits of Agreement (LoA)
RAH vs Q1	-0.12 (-13.39, 13.15)
RAH vs Q2	-0.03 (-15.14, 15.07)
Q1 vs Q2	0.09 (-4.46, 4.63)

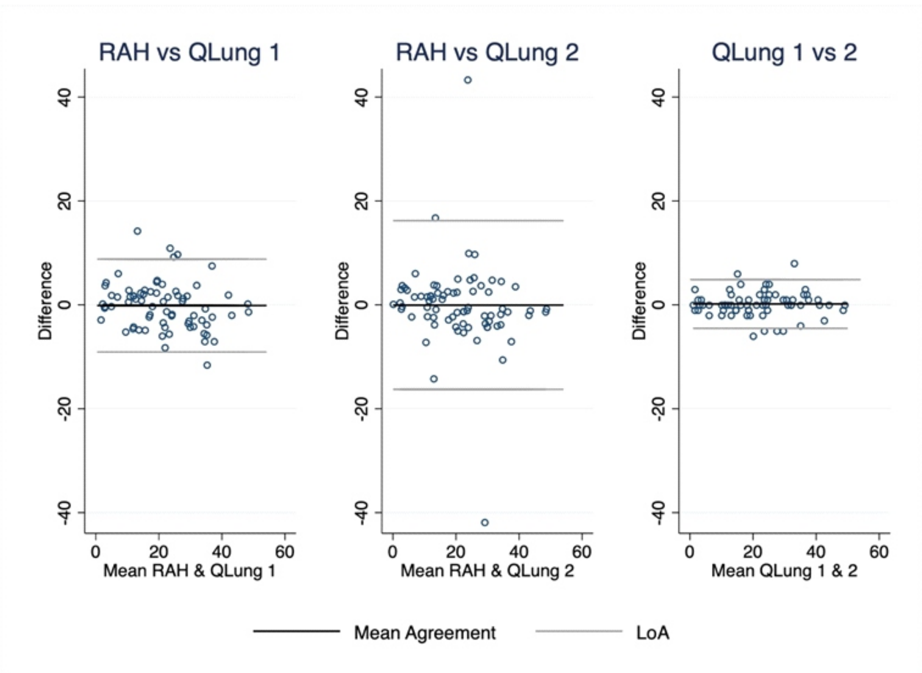
Figure 2c. Lobar differential volume concordance by ICC and by mean difference/limits of agreement.

Table 1. The documented troublesome datasets during Q Lung processing.

Participant	Troublesome Data Encountered	Reason
6	Left lower and left upperlobes for Q1	Q1: Lt oblique fissure
8	Left lower and left upper lobes for Q1 and Q2	Q1: Lt oblique fissure Q2: Lt oblique fissure
12	Left lower and left upperlobes for Q1 Right mid and right upperlobes for both Q1 and Q2	Q1: Lt oblique &Rt horizontal fissures Q2: Rt horizontal fissure
13	Right lower and midlobes for both Q1 and Q2 Right upper lobe for Q2	Q1: Rt oblique fissure Q2: Rt oblique & horizontal fissures
14	Right mid and upperlobes for both Q1 and Q2 Right lower lobe for Q2	Q1: Rt horizontal fissure Q2: Rt oblique & horizontal fissures
15	Right mid and upper lobes for Q1	Q1: Rt horizontal fissure
16	Right mid and right upperlobes for Q2	Q2: Rthorizontal fissure
17	All lobes right lung (lower, mid and upper) for Q1 and Q2	Q1: no visible fissures Q2: no visible fissures
18	Right mid and upper lobes for Q1 and Q2	Q1: Rt horizontal fissure Q2: Rt horizontal fissure
19	Right mid and upper lobes for Q1	Q1: Rt horizontal fissure

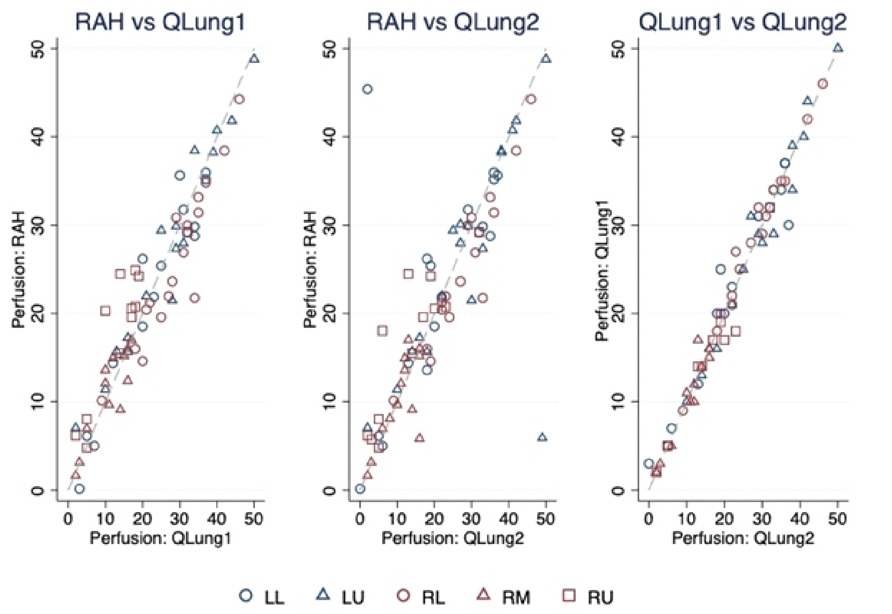


Ventilation		
Comparison	Estimated ICC (95% CI)	Concordance
RAH vs Q1	0.93 (0.88, 0.95)	Very Strong
RAH vs Q2	0.77 (0.67, 0.85)	Strong
Q1 vs Q2	0.98 (0.97, 0.99)	Very Strong

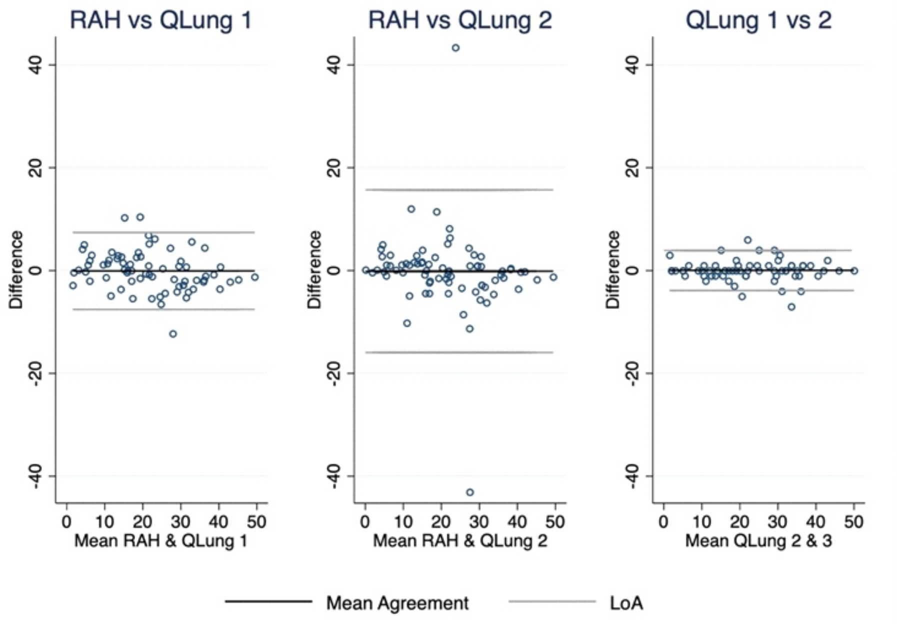


Ventilation	
Comparison	Estimated Mean Difference + Limits of Agreement (LoA)
RAH vs Q1	-0.13 (-9.06, 8.79)
RAH vs Q2	-0.04 (-16.26, 16.19)
Q1 vs Q2	0.18 (-4.53, 4.88)

Figure 3a. Lobar differential ventilation concordance by ICC and by mean difference/limits of agreement, without the troublesome data.

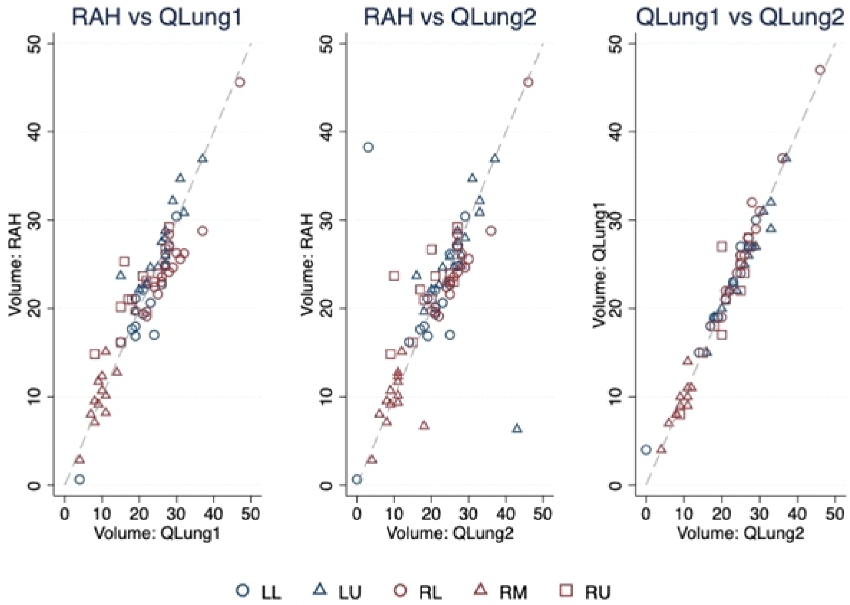


Perfusion		
Comparison	Estimated ICC (95% CI)	Concordance
RAH vs Q1	0.95 (0.91, 0.97)	Very Strong
RAH vs Q2	0.78 (0.67, 0.85)	Strong
Q1 vs Q2	0.99 (0.98, 0.99)	Very Strong

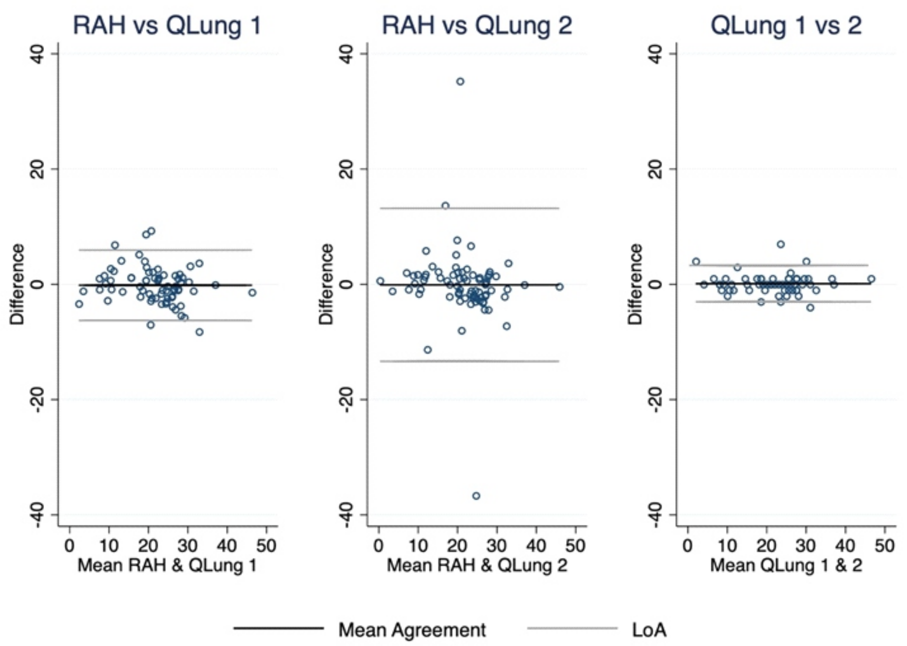


Perfusion	
Comparison	Estimated Mean Difference + Limits of Agreement (LoA)
RAH vs Q1	-0.05 (-7.53, 7.43)
RAH vs Q2	-0.12 (-15.93, 15.68)
Q1 vs Q2	0.03 (-3.86, 3.92)

Figure 3b. Lobar differential perfusion concordance by ICC and by mean difference/limits of agreement, without the troublesome data.



Volume		
Comparison	Estimated ICC (95% CI)	Concordance
RAH vs Q1	0.93 (0.88, 0.95)	Very Strong
RAH vs Q2	0.68 (0.54, 0.79)	Strong
Q1 vs Q2	0.98 (0.97, 0.99)	Very Strong



Volume	
Comparison	Estimated Mean Difference + Limits of Agreement (LoA)
RAH vs Q1	-0.16 (-6.28, 5.96)
RAH vs Q2	-0.10 (-13.36, 13.17)
Q1 vs Q2	0.12 (-3.03, 3.27)

Figure 3c. Lobar differential volume concordance by ICC and by mean difference/limits of agreement, without the troublesome data.

concordance with RAHVQSQ.

One weakness of this study is that the Q lung program used was a modified version that allowed an analysis of VQ SPECT in conjunction with a separate CT. The program is thought to be functionally similar to the commercial version which required the VQ SPECT/CT to be acquired on a GE Discovery or later models of hybrid gamma cameras. We had to resort to this modification as the studies were acquired on an older model of hybrid scanner, the GE Hawkeye 4 and the CT were acquired from separate scanners.

In conclusion, using VQ SPECT/CT data of participants with advanced airways disease, our study has found a close concordance of estimated differential lobar ventilation, perfusion and volume percentages using RAHVQSQ when compared with a duplicated blinded assessment using Q lung which is a commercially available program that is approved as a clinical device. The good concordance supports the validity of our quantitative methodology.

Disclosures

A temporary Q lung licence was provided in kind by GE Healthcare for the purpose of this study without charge. We have no financial interest or relationship with GE Healthcare. GE Healthcare had no input into the design, conduct, analysis or interpretation of this study. The study has been granted approval by the Central Adelaide Health Network human research committee (Reference number HREC/18/RAH/797). Signed informed consent was obtained from participants for the initial study evaluating our technique for bronchoscopic lung volume reduction assessment. The ethics committee waived the need for repeat signed informed consent from the participants for the purpose of this validation study which utilised the same deidentified datasets.

Bibliography

1. Juan C. The Lung Scan in Patient Selection for Lung Volume Reduction Surgery. *J Nucl Med Technol* 1998; 26: 26-9
2. Tee V, Nguyen P, Jersmann H et al. Use of Ventilation-Perfusion Single-Photon Emission Computed Tomography to Select the Target Lobe for Endobronchial Valve Lung Volume Reduction. *Respiration* 2021;100: 886-97.
3. Nijor S, Gangemi A, Patel R et al. Quantitative SPECT/CT Imaging in Pre Op Evaluation of Patients for Bronchoscopic Lung Volume Reduction (BLVR). *J Nucl Med* 2020; 61 (supplement 1): 1493
4. Nicolas G, Chirindel A, Cachovan M et al. Quantitative and anatomical assessment of lung perfusion with ^{99m}Tc-MAA SPECT/CT in emphysema and lung cancer patients. *J Nucl Med* 2018; 59 (supplement 1): 516
5. Chandra D, Lipson D, Hoffman E et al. Perfusion Scintigraphy and Patient Selection for Lung Volume Reduction Surgery. *Am J Respir Crit Care Med* 2010; 182: 937-46.
6. Roach P, Schembri G, Bailey D. V/Q Scanning Using SPECT and SPECT/CT. *J Nucl Med* 2013; 54: 1588-96.
7. Chew C, Hopkins E, Crouch B et al. The Role of VQ SPECT CT Quantification Using RAHVQSQ Program in Pre Lobectomy Assessment for Early Stage Non-Small Cell Lung Cancer: A Pilot Study. *Intern Med J* 2021; 51(S3): 15-38.
8. El-Barhoun EN, Chew G, Crouch B, Chatterton B. Reproducibility of a semi-quantitative lobar pulmonary ventilation and perfusion technique using SPET and CT. *Hell J Nucl Med* 2017; 20(1): 71-5.
9. Carstensen B, Simpson J, Gurrin LJ. Statistical Models for Assessing Agreement in Method Comparison Studies with Replicate Measurements. *Int J Biostat* 2008; 4(1): Article 16.