JTD JOURNAL OF THORACIC DISEASE
A PEER-REVIEWED, OPEN ACCESS JOURNAL FOR HIGH-QUALITY RESEARCH IN THORACIC DISEASES
IMPACT FACTOR 2.895

Peer Review File

Article information: https://dx.doi.org/10.21037/jtd-21-1755

Reviewer A

Comment 1: This manuscript reports on the findings of a small retrospective small clinical study investigating the accuracy of chest X-ray measurement for size matching between donors (n=25) and recipients (n=24) in lung transplantation.

Reply1: A sample size calculation is provided in section 2.2.4 of the methods section.
Changes to the text: none

The authors demonstrated that the accuracy of lung height measurements on chest-x-ray in donors and recipients, and therefore size matching, can be improved with multiple measurements by calculating the mean lung height from several X-rays. This results in better agreement amongst individuals when interpreting lung height for size matching between donor and recipient.

Major Comments:

1) Clinical relevance of chest-x ray measurements versus predicted total lung capacity based on height and gender:
Comment 1: In the reviewers' clinical practice over > 1200 lung transplants, lung height is never measured on chest x-ray and not used for donor to recipient size matching. We rely on predicted total lung capacity of both donor and recipient, calculated according to the formula based on height and gender.

Reply 1: We acknowledge the use of predicted Total Lung Capacity based on height and gender equations as one of the most commonly used methods to size match donors and repients in clinical practice. As we state in the introduction there is no consensus on the best method to size match donors and recipients. As an example Reviewer E uses a combination of predicted TLC and actual TLC for the recipient, we use height, gender and CXR lung height without any particular formulas.

Changes in the text:
A new Table 1 with advantages and limitations of the different size matching methods has been added to the introduction.
"Transplant centers use a variety of methods to size match donors and recipients independently or in combination (Table 1). The most common tools are predictive equations and radiographic estimates of lung size.."

Comment 2: Of course, actual and predicted TLC will differ, but actual TLC (correlating chest x-ray measuremens) is not relevant for somebody waiting for new lungs.The lung transplant candidate is waiting for a "normal/ideal" lung according to his/her height and gender.

Reply 2: We recognize the reviewer's point**.** Unfortunately**,** height and gender are not accurate predictors of TLC at an individual level. The equations are designed to differentiate healthy from diseased individuals and the variability between healthy individuals is very large. The average healthy male in the United States is 179 cm tall, his predicted TLC ranges between 5.3 and 7.7 liters. Therefore, it is not an ideal tool to ascertain the "ideal" lung. What is the maximum or minimal predicted TLC ratio acceptable for a particular donor-recipient pair? pTLC equations only allow us to make gross approximations, but we can't answer this question. This is one of the reasons why size mistmatch continues to occur and results in adverse outcomes for patients.

We argue that more precise measuring methods are needed to prevent size mismatch.

Changes to the text:
The following changes have been made in the introduction:
"…Practice patterns indicate that "real" lung size matters with systematic acceptance of larger donor lungs for recipients with chronic obstructive pulmonary disease and cystic fibrosis and smaller donor lungs for recipients with interstitial lung disease (8)."

"The use of predictive total lung capacity (TLC) equations has the theoretical advantage of providing TLC estimates unbiased by the underlying disease of the recipient or the use of mechanical ventilation in the donor. Predicted TLC equations are meant to differentiate healthy and diseased individuals and were not designed to be precise at the individual level (9). For this reason, they have wide confidence intervals. As an example, an average healthy male in the United States is 179 cm tall, and his predicted TLC ranges between 5.3 and 7.7 liters (10–12). This lack of precision is a barrier to establishing the acceptable limits of size discrepancy."

Comment 3: The chest x-ray in the donor may not reflect the real height of the donor lung because of possible atelectasis as a result of underventilation with the donor in a supine position in the ICU.

Reply 3: We agree with the reviewer that atelectasis, pneumonia, pleural effusions, radiographic technique including issues with magnification, rotation can all affect measurements. In these situations, the physicians provide their best estimate. Over multiple measurements, the lung heights are approximately normally distributed and

converge around the mean. All these sources of error support the need for multiple measurements.

Changes to the text:
We have added the number of CXRs with infiltrates to the results section "For the rest of the cohort, 29 (7.5%) CXRs had basilar infiltrates/effusions."
We have also updated the methods section "In cases were basilar infiltrates obscured the diaphragm or costophrenic angles, the raters used their best judgement to provide measurements."

Comment 4: Equally, the chest-x ray in the recipient does not reflect the "normal" (predicted) total lung capacity as the lung can be hyperinflated in a COPD patient or restricted in a fibrotic patient. Matching should be based on the ideal lung height to be implanted in a particular recipient. Lung height measured on chest x-ray in a lung transplant candidate will only reflect the actual lung size, but not the preferred lung size.
So what is the clinical relevance of measuring lung height on chest X-ray in a transplant candidate? In my opinion, this is useless.

Reply 4: We thank the reviewer for his comment and we respectfully disagree**.**
The reviewer states that the "preferred" lung size for recipients is their predicted TLC. The problem with this statement is the lack of precision of TLC estimates (see reply 2). There's not one predicted TLC, there's a range of predicted TLC and the range is very wide.
In our article we do not suggest matching donors to recipients based on equal lung heights. The preferred lung height difference between donor and recipient for each underlying diagnosis will need to be the subject of future studies.

Changes to the text:
The following paragraph has been modified in the discussion:
"These results highlight the association of predicted TLC and CXR lung heights with outcomes at the population level. At the individual level they both fail to establish accurate size relationships due to their lack of precision. Predictive TLC equations consider sex and age but TLC (12) is dependent on many other factors such as genetic and epigenetic influences, preterm birth, early childhood infections, malnutrition and other noxious exposures (25–27). These factors can't currently be accounted for reliably and are responsible for the predictions' wide confidence intervals. The precision of these equations has not improved over the past 50 years and it should not be expected to improve (12). Predictive TLC equations were conceived to differentiate disease states from health, not to be precise. They allow the gross description of relationships between size and outcomes but may not the right tool for personalized size matching."

The following sentence has also been added to the discussion:
"For this reason, we do not describe an ideal lung height difference between donors and recipients.. Future studies using the mean CXR lung height are needed to answer this question."

Comment 5: The authors should better explain in the discussion the clinical relevance of size matching based on chest X-ray in transplant candidates with end-stage lung disease. The statement in the introduction that "Chest x-Ray (CXR) linear measurements are the most common method to estimate "real" TLC during donor-recipient matching" does not fit with the current clinical practice in many lung transplant programmes across the world.

Reply 5: We acknowledge the reviewer's point. It is unclear what the most common method is. The 2019 International Society for Heart and Lung Transplantation annual report discusses clinical practice in size matching: the mean height difference between donors and recipients was + 5 cm for COPD, -1.9 cm for ILD recipients and +4.1 cm for CF. We wouldn't expect to observe these differences if pTLC was used uniformly without attention to actual TLC.

Changes to the text:

"Practice patterns indicate that "real" lung size matters with systematic acceptance of larger donor lungs for recipients with chronic obstructive pulmonary disease and cystic fibrosis and smaller donor lungs for recipients with interstitial lung disease (8)."

"Average lung height measurements across multiple CXRs are more precise and should be validated against clinical outcomes before their implementation in clinical practice."

Minor Comments

Comment 3: 2) Table 1(now table 2):
As listed in Table 1, right lung height (20 cm) was about 2 cm smaller than left lung height (22 cm), both in donor and recipient. How to explain this difference knowing that the right lung is bigger in size (10 segments) compared to left lung (9 segments).

Reply 3: Both lungs are three-dimensional structures, the relationship of lung height to volume is different for the right and left lung. Using the volume formula for a truncated cone it is easy to see how a shorter structure can have a larger volume. Let V be the volume, r_1 the radius of the upper circumference and r_2 the radius of the base

circumference and h the lung height:

$V= 1/3\pi(r_1^2 + r_1r_2 + r_2^2)h$
If h= 22 for the left lung, r_1=5 and r_2=9, then V= 3.5 liters
If h= 20 for the right lung, r_1=5 and r_2=10, then V= 3.7 liters

Changes in the text:
The following text has been added to the discussion:
"Finally, the CXR lung heights are only measuring one dimension, but the lungs are three-dimensional structures. This results in the apparent incongruency of the right lung being shorter on average than the left lung (20 vs 22 cm) even though its volume is usually larger. A shorter but wider or/and deeper structure can have a larger volume and the relationship between lung height and volume is different for the right and left lungs for this reason. Quantitative computed tomography could overcome this limitation. However, in clinical practice it will likely be subject to errors in measurement similar to those observed in our study (28)."

Reviewer B
Comment 1: Nice initiative of these authors. The topic matter concerns an interesting subject, however requires several corrections. Also, the topic matter of this original article will be of very narrow interest to the readership of this journal. As it stands this article is perhaps better suited to a journal solely focused on transplant field.

Reply1: Thank you for your review. Although our article is focused on lung transplantation, estimation of lung volumes from radiographic measurements is a topic of general interest in thoracic surgery and pulmonology. Error measurement analysis, widely used amongst ingeneers could be applied to a variety of topics in thoracic surgery: e.g. how does error in measurement impact performance of lung cancer survival predictive scores? We hope both aspects of our article will be of interest to researchers in thoracic surgery.

Changes to the Text:
None

Reviewer C
The authors of this interesting manuscript investigate the precision and accurancy of CXR in eveluating donor-recipient mismatch after lung transplantation. The authors concluded that not the single but the averaged CXR measurements may be helpful in estimating donor-recipient mismatch. I have the following comments:

Comment 1: Page 4, line 117: please replace unbiassed with unbiased.
Reply1: We have done the required change.

Changes to the text: Unbiassed has been changed to unbiased.

Comment 2: Page 5, line 139: it is not well specified, if the 25 selected donors are the donors of the included 25 recipients.
Reply2: Thank you for pointing this out. They are not. The focus of the study is the precision of lung height measurements and their impact on size matching donors and recipients, not outcomes of size mismatched procedures. The text has been clarified.
Changes to the text:
"…and 25 donors from all prospective brain-dead donors managed at MAT's independent organ recovery center over the same period were included. Donors and recipients were independent of each other.."

Comment 3: Page 5, line 153: it should be written "...Hospital and (not And) 25 donors...".
Reply 3: The requested changes have been made.
Changes to the text: The text now reads: ".. and 25 donors…"

Comment 4: Page 6, lines 173-174: in how many cases the dome of the diaphragm was not apparent? and Why? How did the authors, for example, account for small pleural effusions that may make the identification of the costophrenic angle difficult?

Reply 4: We acknowledge the reviewer's question. The focus of the study is the evaluation of the precision of CXR lung height measurements in real life practice. In the cases where effusions or infiltrates obscured the diaphragm we requested raters to provide their best estimate. This instance was unusual and the one patient where all CXRs had bilateral lower lobe infiltrates was excluded from analysis. The rationale for exclusion was that his measurements resulted in clear outliers with differences in lung height measurements > 3 sd compared to the rest.

Changes to the tex:
"In cases were basilar infiltrates obscured the diaphragm or costophrenic angles, the raters used their best judgement to provide measurements."

"One recipient had bibasilar radiographic infiltrates obscuring both diaphragms in each of the 6 chest x-rays resulting in differences in height greater than 3 times the SD and was excluded from the analyses. For the rest of the cohort, 29 (7.5%) CXRs had basilar infiltrates/effusions."

Comments 5: Page 8, lines 223-225: see point 4): why did the authors then exclude this patient from the analysis?

Reply 5: The rationale for excluding him was that his measurements resulted in clear

outliers with differences in measurements > 3 sd compared to the rest.

Changes to the text:
"One recipient had bibasilar radiographic infiltrates obscuring both diaphragms in each of the 6 chest x-rays resulting in differences in height greater than 3 times the SD and was excluded from the analyses. For the rest of the cohort, 29 (7.5%) CXRs had basilar infiltrates/effusions."

Comment 6 : Page 10, line 282: the authors may discuss the possible role of chest computed tomography instead of CXR in evaluating donor and recipient size mismatch. For example, 3D Imaging may be possible with CT and not CXR.

Reply 6: We thank the reviewer for broadening the discussion**.** Although 3D imaging is promising and has been used successfully for size matching in living lobar lung transplants, CT imaging is likely to be subject to the same measurement errors as CXR: e.g. respiratory phase timing will bring an error in measurement that will have to be accounted for. Our study illustrates this shortcoming of radiographic methods to size match donors and recipients in lung transplantation and could guide future research using CT volumetry. We have added a comment to the discussion and a new reference that illustrates this shortcoming of 3D imaging.

Changes to the text:
"Finally, the CXR lung heights are only measuring one dimension, but the lungs are three-dimensional structures. This results in the apparent incongruency of the right lung being shorter on average than the left lung (20 vs 22 cm) even though its volume is usually larger. A shorter but wider or/and deeper structure can have a larger volume and the relationship between lung height and volume is different for the right and left lungs for this reason. Quantitative computed tomography could overcome this limitation. However, in clinical practice it will likely be subject to errors in measurement similar to those observed in our study (28)."

Reviewer D

This is very nice in-depth manuscript showing that averaging multiple chest X rays reduces interobserver variability and therefore would theoretically improve the precision of size matching for lung transplantation. The manuscript further showed that the imprecision of using single CXR could lead to missed transplant opportunities, which could be avoided with multiple CXRs. This would be a simple, safe, and low-cost practice to implement. The main limitation toward its use would be the time required to measure multiple CXRs.

Comment 1: The authors state that there were a median 4 CXRs per subject (IQR

3-6). As a a practical matter, it would be helpful to know if subjects with more CXRs had less error. i.e., how many CXRs must one measure (or how many must be available) for this method to be effective? Was this included in the initial univariate linear regression analysis (line 232)?

Reply 1: We thank the reviewer for his comments. In this manuscript**,** we rellied on our data as an example of how use of the existing CXRs can be maximized. It is posible to calculate the number of CXRs needed to reduce error to the desired precision. The formula is as follows:

N =(1.96 * sd/ desired margin of error)

Changes to the text: "The average number of CXRs in our study was 4 and this was enough to achieve a margin of error for the mean lung height of approximately 1.5 cm. More measurements would decrease this margin of error according to the formula N = (1.96*SD)/desired margin of error."

Comment2 : The authors assert that an average lung measurement (arithmetic mean) is appropriate since the random error takes a normal distribution. This was well demonstrated in the supplementary materials. In practice, I think clinicians would be prone to remove outlier values for lung height and average the remaining CXRs that fall in close proximity. Was there any investigation into whether median (rather than mean) CXR height would give similar results?

Reply 2: We recognize the reviewer's point. Since the data approximated the normnal distribution, the mean and median would be expected to be very similar. Based on our results we would recommend including all available data in the calculation and not excluding presumed outliers.

Changes to the text: none

Comment 3: Typos noted: Line 340 (steaming should be stemming), Line 304 (ration should be ratio), Line 218 (alfa should be alpha).

Reply 3: the requested changes have been made to the text.

Changes to the text: steaming, alfa and ration have been changed to stemming, alpha, ratio

Reviewer E

This manuscript describes a retrospective series looking at the accuracy of chest x-ray

measurements for size matching in lung transplantation. Measuring chest rays to perform size matching in lung transplantation does not represent a state of the art method since superior matching methods based on total lung capacity, which is based on height and gender, are available.

Reply 1: We thank the reviewer for his comments and we acknowledge the use of predicted Total Lung Capacity based on height and gender equations as one of the methods used to size match donors and repients in clinical practice. As we state in the introduction there is no consensus on the best method to size match donors and recipients. As an example Reviewer A uses predicted TLC for donor and recipient, we use height, gender and CXR lung height without any particular formulas and the reviewer reports a mix of predicted TLC in the donor and a combination of actual TLC from pulmonary function tests and predicted TLC in the recipients.

Changes to the text:
A new Table 1 with advantages and limitations of different size matching methods has been added to the introduction.

Comment 2: Chest x-rays in the ventilated donor in supine position do not necessarily reflect true lung size

Reply 2: We agree with the reviewer in stating that the lung height of a supine patient does not necessarily reflect the "maximal" lung height, this is also true for the recipient. One could choose the CXR with the largest lung height available, but wouldn't be able to tell if it was the largest possible. The mean lung height overcomes this limitation. In our analysis the range of error for lung heights (in both donors and recipients) was approx. 6cm, similar to previous studies with double exposure CXRs at full inspiration and exhalation. Also the largest contributor to error measurement was respiratory variability between images. This 2 findings suggests that the average lung height should be comparable between donors and recipients.

Changes to the text:
"Previous studies using double exposure CXR (full inspiration and expiration) found an average diaphragmatic excursion of around 6 cm (14), very close to the range of CXR lung height error in our analysis"

Comment 3: …and in the recipient calculation of the optimal donor lung size by taking predicted and real TLC into account is much more accurate and no additional value is gained by measuring chest rays.

Reply 2: We acknowledge the reviewer's point and we would like to refer to reply 2 for reviewer A. In brief, the predicted TLC is not a single value, it is a range. In the

case of the average American male, 179 cm tall, the predicted TLC will vary between 5.3 to 7.7 liters. This will create a large uncertainty for the donor and the recipient even after adjusting for the actual TLC.

Changes to the text:

"The use of predictive total lung capacity (TLC) equations has the theoretical advantage of providing TLC estimates unbiased by the underlying disease of the recipient or the use of mechanical ventilation in the donor. Predicted TLC equations are meant to differentiate healthy and diseased individuals and were not designed to be precise at the individual level (9). For this reason, they have wide confidence intervals. As an example, an average healthy male in the United States is 179 cm tall, and his predicted TLC ranges between 5.3 and 7.7 liters (10–12). This lack of precision is a barrier to establishing the acceptable limits of size discrepancy."

Reviewer F

Coment 1:

The reviewer is honored to review an article about size matching using chest X-ray films. It is really an important issue in lung transplantation, but has not been well discussed. This paper studied chest X-ray lung height measurements and size matching by using chest X-ray films of the donor and the recipient. The authors did not discuss either early and late outcomes.

Reply 1: We thank the reviewer for his comments. We agree, it will be interesting to learn about the impact of using average CXR lung height in short and long term outcomes. Our study represents a preliminary step to such analysis. We wanted to assess the "tool's" performance, its precision and repeatability. We have attempted to clarify the methods.

Changes to the text:

"Donors and recipients were independent of each other."

Coment 2: In this point, scientific meaning of this paper is limited, but the size matching using chest X-ray films is much easier than that using chest CT.

Reply 2: We agree with your observation, clinical CT scans will likely suffer from the same intraindividual variability. The information learned in this study will inform future studies aiming to use CT scan volumetry for size matching in lung transplantation.

Changes to the text: "Finally, the CXR lung heights are only measuring one dimension, but the lungs are three-dimensional structures. This results in the apparent

incongruency of the right lung being shorter on average than the left lung (20 vs 22 cm) even though its volume is usually larger. A shorter but wider or/and deeper structure can have a larger volume and the relationship between lung height and volume is different for the right and left lungs for this reason. Quantitative computed tomography could overcome this limitation. However, in clinical practice it will likely be subject to errors in measurement similar to those observed in our study (28).

Coment 3: The reviewer wants to know a clinical difference in size matching between chest X-ray and pulmonary function test (FVC and/or TLC). It might take pains for the authors to do such comparison, but it might possibly provide some meaningful results.

Reply 3: We acknowledge the reviewer's question. Unfortunately the reason we need to use estimates of lung volumes is the impracticality of obtaining pulmonary function tests in the brain dead mechanically ventilated donors.

Changes to text: we have added a table with methods used to size match donors and recipients and some of their limitations. We have included the inability to perform PFTs in the donor as one of the limitations Table 1.

Comment 4: There are many typos as well as grammatical errors in this manuscript, so please ask a professional English editor who is familiar with scientific writing.

Reply 4: The manuscript has been edited as requested.

Comment 5: Further, statistical review should be performed by a professional statistician.
Reply 5: The author team includes Michael Wallendorf, PhD a senior research statistician from the Division of Statistics at Washington University School of Medicine.

Minor points:

1) There are many abbreviations in this manuscript. Please spell them out at the first use.

Reply: the text has reviewed as requested

2) There are several words and phrases with an underline or in an italic type. Please fix them.

Reply: The text has been reviewed as requested

3) On line 241, "8cm" should be "8 cm". There are many typos of this kind throughout the manuscript.

Reply: the text has reviewed as requested

4) Please check references again.

JOURNAL CITATION REPORTS® IMPACT FACTOR 2.895 ©2021 Clarivate Analytics

Reply: the references have been reviewed for accuracy and format.