

Peer Review File

Article information: <https://dx.doi.org/10.21037/atm-21-5069>

Reviewer A:

I think the task and importance of this paper is very high.

I am using 3D printing in daily routine (low cost printer + PLA + Ultimaker) and it shows, as you have shown in your study the most important information before surgery. So I think the result and conclusion from your study, that even a low-cost 3D printer may be relevant for preoperative planning, could be used to get the funding from a hospital board for this very helpful tool. Therefore, your cost calculation is very relevant for this discussion.

Reply: We appreciate the reviewer's acknowledgement on the importance of this work.

Although your study focusses on such a very important and relevant task, there is a huge problem concerning the results of your survey.

First of all - why did you use such a small group? 5 medical professionals and only ONE orthopedic surgeon. As you showed in your results, most of the question in your survey weren't answered, although they should have.

Why would the anesthesiologist in your study now how to treat a acetabular fracture or in most cases how to classify this fracture pattern. Of course you could provide a specialist and a control group, but I would recommend to repeat this survey.

Reply: We plan to expand our limitation section of this study to address these concerns. Unfortunately, at this time we do not have the funding to repeat the survey or to include a larger sample. We hope to improve our recruitment and sample size in future studies and will address these concerns more thoroughly in the limitation section of the current manuscript. The additional statement has been included in line 281 - 286:

Changes in the text: Page, 13, Line 281 – 286: “However, one major limitation from this study is the quality of information that can be drawn from the survey portion. The models were reviewed by 5 participants in different fields of medicine, not all of which would use the model for the intended surgical planning purpose and may be subject to a fair amount of bias. The survey suggests that the models are representative and accurate, however may not be conclusive from this study alone.”

Also it has to be clear that the differences in measurement between the models seem to be significant, but we are talking about .4 - .6 mm, I think there is no relevance in preoperative planning which such a difference, only when a 3D printed wire guide is used intraoperatively. So this has to be answered by more than 1 surgeon...

Reply: In order to make the distinction of the small differences between the models, we have added the following statement to the discussion:

Changes in the text: Page 14, Line 306 - “However, the error differences found in this study were extremely small, often times less than 1mm, and would likely not play a significant role in surgical planning for orthopedic purposes.”

Besides this there are some smaller writing and editing mistakes which should be corrected.

Reply: The manuscript has been reviewed additional times to ensure writing and editing mistakes have been addressed.

Reviewer B:

Quite straightforward study, comparing a low-cost set-up with a high-cost set-up for the evaluation of 3D printed acetabular fracture cases.

It is nice to see whether a low-cost pathway is non-inferior to a high-cost pathway.

Reply: Once again, we appreciate the reviewer's acknowledgement on the importance of this work.

Legislation is missing from the discussion. Is it legally possible to use open source information to plan surgical procedures?

Reply: Regulation and commercialization are difficult to control in open-source applications. Currently, 3D Slicer is not approved for clinical use and is intended for research purposes, which does not impose restrictions on its usage. However, that does not mean it cannot still offer useful information prior to a surgical procedure, especially in more complex cases. This distinction has been added to the discussion at line 271:

Changes in the text: Page 12, Line 271 - 274: "Currently, 3D Slicer is intended for research purposes, rather than clinical application like Mimics is. However, the results of this study indicate that the open-source software could offer some utility to institutions that do not have the resources for proprietary software."

Time is missing as an outcome, it would be interesting to know how long segmentation needed, planning and post-processing. Time is money. Therefore it is necessary to take into account if you want to make cost-benefit analyses.

Reply: Time is a very important step in the segmentation procedure, and can greatly contribute to its cost. Segmentation time can vary greatly depending on case complexity, personnel experience, and software familiarity. In order to address these points, we will be adding the following statements to the manuscript:

Changes in the text: Page 5, Line 111 - 114: "Segmentation was performed using a combination of semi-automatic and manual thresholding techniques by experienced medical image segmentation specialists and verified by a licensed radiologist. Each segmentation on both 3D Slicer and Mimics took a similar amount of time, at approximately 60 minutes per model."

Page 6, Line 116 – 120: “Post-processing procedures varied per printer: Objet model support was removed via a high-pressure water cleaner (Powerblast, Balco, UK), uPrint support was removed via a dissolvable support solution (ABS, Stratasys, Inc. Eden Prairie, MN, USA), Form 2 support was removed manually and soaked in an isopropyl alcohol solvent bath, and Ultimaker support was removed manually.”

Additionally, we have included a new table (now Table 1) to provide additional information about the average printing times.

Methods:

- It not evident that the order in which the method section is being presented is also the order in which all measurements were performed. Was there blinding of the observers? Did they do the measurements on the 3D models first or first the measurements on the CT etc.

Reply: CT measurements and physical measurements were taken separately to prevent any form of bias. Additionally, the individual taking the measurements was blind to which model they were measuring at the time as an additional measure of bias prevention. The following change has been made to the manuscript:

Changes in the text: Page 7, Line 158 – 163: “First, digital measurements were taken using Change PACS. The measurements were then repeated on the 3D renderings after segmentation using the respective segmentation software. Then, separately, physical measurements taken from the 3D printed models were recorded using high-grade digital calipers (PEC Tools, Santa Monica, CA, USA) by a second observer who was blinded to the model type (15).”

- Which software is being used to perform the CT measurements? Give more information about the CT measurements as this is your “ground truth”.

Reply: CT Measurements were calculated using Change picture archiving and communication system (PACS). (Change Healthcare, USA). The following sentence was added to line 149 to indicate this change:

Changes in the text: Page 7, Line 149 – 152: “In order to measure model accuracy in displaying the region of interest, digital measurements were taken on the original CT

imaging data using the Change picture archiving and communication system (PACS) (Change Healthcare, USA) of the distance between specific landmarks and then replicated on the physical 3D printed models.”

- Is the segmentation checked by a second observer? What if manual mistakes were made? How were those checked?

Reply: Segmentations were performed by two of the co-authors and verified by a certified radiologist, all of whom are experienced with cross-sectional anatomy and image segmentation for medical 3D printing. The following line has been added to make this clear to the reader. This was addressed in a previous comment but can be found in line 111:

Changes in the text: Page 5, Line 111 – 113: “Segmentation was performed using a combination of semi-automatic and manual thresholding techniques by experienced medical image segmentation specialists and verified by a licensed radiologist.”

- How were the models printed? Does the direction of printing have influence on non-resorbable support removal or X-Y-Z accuracy of the models?

Reply: To standardize the layering of each of the printing methods, all models were printed in the same orientation, with the acetabulum facing upward on the build platform. A statement discussing support removal was added at line 116. *The support question mentioned is discussed in the next comment.* To address the comment about the print orientation we have added line 114 to the manuscript:

Changes in the text: Page 5, Line 114 – 116: “To standardize the orientation of each of the printing methods, all models were printed with the acetabulum facing upward on the build platform.”

- What was done with support material, removal etc.? Could this have affected measurements?

Reply: Support was placed and removed in locations that did not contact the measured locations and would not affect the model’s build quality and subsequent measurements. A statement discussing support removal was added at line 116.

Changes in the text: Page 6, Line 116 – 120: “Post-processing procedures varied per printer: Objet model support was removed via a high-pressure water cleaner (Powerblast, Balco, UK), uPrint support was removed via a dissolvable support solution (ABS, Stratasys, Inc. Eden Prairie, MN, USA), Form 2 support was removed manually and soaked in an isopropyl alcohol solvent bath, and Ultimaker support was removed manually.”

- Manual segmentation error vs. Printing error, measurements on segmentation?

Reply: We recognize that the segmentation procedure is a large potential source of error. We would like to thank the reviewer in their recommendation to perform this analysis and have added this to the manuscript. The relevant information has been added to line 204:

Changes in the text: Page 9, Line 204 - 208: “Measurements of the distance between the PIIS to the ischial spine and the acetabular fracture fragment were assessed with the respective software that was used for the segmentation process and can be seen in Table 2. The mean absolute error from the software measurements was found not to be statistically significant between software types for both the PIIS to ischial spine distance and the fracture fragment distance, $F(3,19) = 0.426$, $p = 0.737$.”

- No cross-over between different methods. The authors could have compared different segmentations method with each other. Therefore errors that are introduced in the segmentation could influence the post-production measurements.

Reply: This comment is appropriately addressed in the previous statement and can be found at line 204.

Results:

- Clinical survey analysis, This section is poor as there is no systematically gathered information of all observers. This may result in reporting / selecting bias.

Reply: The potential bias of the survey has been addressed as a limitation within the discussion, and can be found on line 281:

Changes in the text: Page 13, Line 281 – 286: “However, one major limitation from this study is the quality of information that can be drawn from the survey portion. The models were

reviewed by 5 participants in different fields of medicine, not all of which would use the model for the intended surgical planning purpose and may be subject to a fair amount of bias. The survey suggests that the models are representative and accurate, however may not be conclusive from this study alone.”

- N=2 for the survey has not enough power to give a significance difference in ‘quality of information the models provided’. Therefore the statement in the results section should be removed / changed.

Reply: We have modified the statement in the results section to clearly indicate that the survey has not enough power to give significant different in ‘quality of information the models provided’. We have made sure to address this with the section added in the previous comment, as well as added the following statement to line 240 of the manuscript:

Changes in the text: Page 11, Line 240 – 242: “The low sample size of the survey and number of responses indicate the survey is subject to potential bias and will be addressed more thoroughly in the discussion.”

Discussion:

- In the third paragraph of the discussion the results from the survey are somewhat being presented as objectified data. However there is likely bias by the observer and asking him the question whether the model is an accurate representation of the general anatomy has a high likelihood of being answered positive. Especially since some questions are only answered by N=2. I suggest a thorough review of this section.

Reply: We have modified the statement in the discussion section to clearly indicate the potential bias. We will also include this as a part of our limitations.

- In the costs paragraph only the costs of material are being examined. Not counting the hours of segmentation by a technician, planning, post-processing of the print, acquisition costs for 3D printer etc. Review this section.

Reply: Additional information about the costs associated with segmentation time, printing time, and post-processing have been added to the manuscript. Table 1 has been added to provide information regarding the printing times.

Changes in the text: Page 5, Line 111 – 120: “Segmentation was performed using a combination of semi-automatic and manual thresholding techniques by experienced medical image segmentation specialists and verified by a licensed radiologist. Each segmentation on both 3D Slicer and Mimics took a similar amount of time, at approximately 60 minutes per model. To standardize the orientation of each of the printing methods, all models were printed with the acetabulum facing upward on the build platform. Post-processing procedures varied per printer: Objet model support was removed via a high-pressure water cleaner (Powerblast, Balco, UK), uPrint support was removed via a dissolvable support solution (ABS, Stratasys, Inc. Eden Prairie, MN, USA), Form 2 support was removed manually and soaked in an isopropyl alcohol solvent bath, and Ultimaker support was removed manually.”

- Line 282 – 299, The question whether segmentation or the printing was the main error in the process is not yet eliminated. It would have been wise for the authors to do a measurement on the segmentation of the 3D model BEFORE 3d printing and after 3D printing. That way the error of 3D printing and segmentation could have been isolated. Now the statement on Line 383 ‘This study was designed to test the entire production method from segmentation to final product with this application in mind is not valid. The author did not consider segmentation as an error.

Reply: We have added this analysis to the manuscript and can be seen in Table 2, as well as added a statement to line 204 regarding the potential error introduced in the segmentation process:

Changes in the text: Page 9, Line 204 – 208: “Measurements of the distance between the PIIS to the ischial spine and the acetabular fracture fragment were assessed with the respective software that was used for the segmentation process and can be seen in Table 2. The mean absolute error from the software measurements was found not to be statistically significant between software types for both the PIIS to ischial spine distance and the fracture fragment distance, $F(3,19) = 0.426$, $p = 0.737$.”

Figures:

- Table 1: Add question numbers to each column

Reply: The questions numbers have been added to each of the columns in the respective Table.

- Figure 5: no intra of inter observer limits of agreement are being researched. These are especially informative in a bland Altman analysis.

Reply: We agree that the observer limits are important and will add it for future studies investigating these technologies. The lack of proportional bias in combination with the non-significant linear trends have previously been reported in literature and should shed light on the acceptable limits of the measurements.

Minor comments:

- Line 49: ‘10% of all pelvis related injuries’, how many overall pelvis injuries are there then?

Reply: This sentence has been changed in the document and can be found on line 48:

Changes in the text: Page 3, Line 48 – 50: “Acetabular fracture is defined as damage to the pelvic bone at the hip joint, and accounts for roughly 10% of the 300,000 hip fractures that occur in the US every year; with an average 37 per 100,000 of these causing damage to the acetabulum as well (1,2).”

- Line 54: ‘High-impact impact incidents’?

Reply: The repeated ‘impact’ has been removed, and the sentence now reads as:

Changes in the text: Page 3, Line 54 - 57: “The necessity for surgical intervention has been most commonly used in high-impact incidents (such as motor incidents), which have a higher occurrence in developing countries due to things such as industrialization, less access to healthcare, and an increased number of motor vehicles per capita (2,3).”

- Line 54: Specify developing countries. Do developing countries have a higher number of cars per capita than developed countries? Or is it motorized vehicles?

Reply: We have added more information to the developed countries portion and redefined cars to motor vehicles and can be seen in the previous comment at the same line (54).

- Line 107: Which mimics version?

Reply: Version information added, the section now reads:

Changes in the text: Page 5, Line 109 – 110: “...Mimics (Version 20.0, Materialise, Belgium; cost \$12,975.00)...”

- Line 108: Which 3D slicer version?

Reply: Version information added, the section now reads:

Changes in the text: Page 5, Line 110 – 111: “...and 3D Slicer (Version 4.8.1, Brigham and Women’s Hospital, Boston, MA; cost free, open-source)...”

- Line 134: Why 10 mm reformats? Why not 1 mm reformats?

Reply: 10 mm reformats were sufficient to identify the landmarks of interest on the multiplanar reformation images. To address this, we have added the following to the manuscript on line 152:

Changes in the text: Page 7, Line 152 – 153: “Measurements were performed on multiplanar maximum intensity projection (MIP) 10 mm reformats, which were sufficient to identify the landmarks of interest.”

- Lines 153-164: If needed move this section to the supplementary data.

Reply: The list of survey questions was removed from the main document and placed as supplementary material 1 (S1).

- Line 182: this needs to move to the methods section

Reply: Model price information was moved to the methods section and can now be found on line 138.

- Line 183-186, incorporate these numbers in one of the tables. It would be nice to see the costs aligned with the accuracy.

Reply: An additional table (now Table 1) was added to display these values in a more convenient form. The following sentence was added to the document.

Changes in the text: Page 7, Line 144 – 145: “Additional information regarding printing time can be found on Table 1.”