Editor and Reviewer comments:      
  
  
  
AE1.0:  
  
**also consider the applicability of this method in open low-field MRI systems which offer an extended motion range.**   
  
RE1.1:  
**Keywords: The first used keywords are already part of the title. To improve SEO, I suggest replacing them with further information about the sequence (i.e. trajectory?) or algorithms used in segmentation**

We thank the reviewer for this suggestion to improve SEO. We have replaced the previous keywords with these: Radial golden-angle acquisition; Canny edge detection; CINE reconstruction; Rigid body tracking

Materials and Methods  
RE1.2: **I would prefer a photograph of a subject in the motion device. I don't find the pictures in [25] to be very helpful. I suggest an image where the patient table is in its home position, to better show the deflection of the lower leg.**

RE1.3:  
**To estimate the flexion-speed better, I suggest to report the cycle time, rather than the cycles per minute. This way it is easier to relate motion speed and sequence temporal resolution.**  
We have revised the text to report cycle time instead of cycles per minute for better clarity. The text on Page 5 now reads: “ Each knee extension-flexion movement cycle was guided by eight metronome beats, with the knee fully flexed at the first beat, fully extended by the fourth beat and fully flexed again by the eighth beat, equating to 8 seconds per cycle”

RE1.4:  
**TGV was used as a regularizer. Along which dimensions? time, space, or both? Please clarify.**  
The TGV regularization was applied along the spatial dimension. The text on Page 6 now reads “... the image was reconstructed as a spatially regularized Total Generalized Variation ~~regularized~~ least-squares problem solved using the Alternating Direction Method of Multipliers, ...”

RE1.5:  
**Is there a python image-processing package that was mainly used for the analysis (i.e. openCV)? If so, please mention the used package.**

SciPy and sciki-image were used for the analysis. We have now mentioned the specific modules used for edge detection, connected-component labelling, cubic spline interpolation and Nelder-Mead optimization alongside the corresponding algorithms in Pages 6 and 7.

RE1.6:  
**In (I) you write: "…including the interior cortical bone boundaries." I suggest: "including the boundary between cortical and trabecular bone"**

We thank the reviewer for the suggestion. The text has been edited accordingly in Page 6.

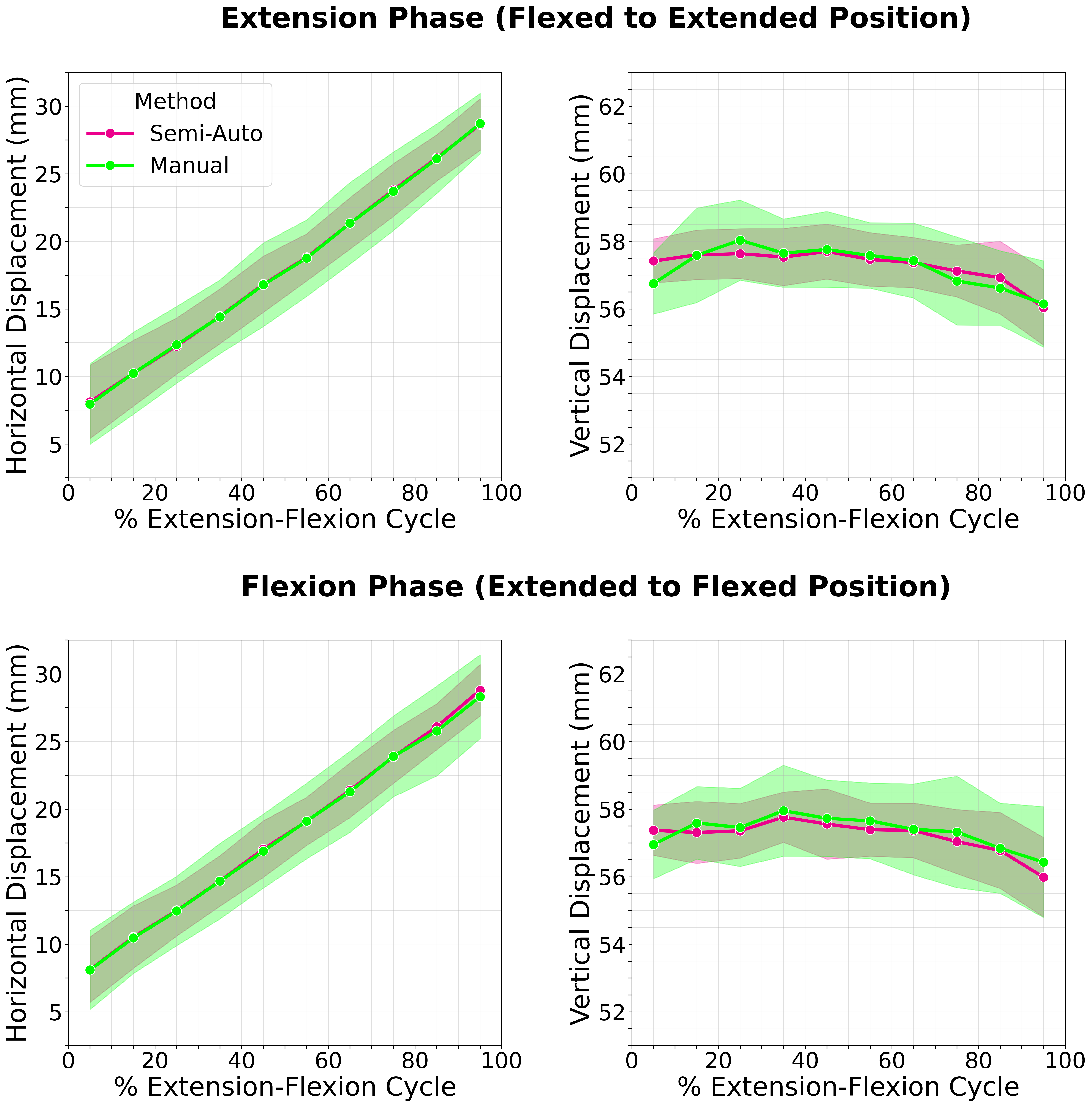
RE1.7:  
**In (III) you write that the most distal points were used as initial points. Would it not be the most proximal point for the tibia?**  
The most distal point was used for the tibia for this study. However, the algorithm can be used with the proximal point as well. The same applies to the femur.

RE1.8:  
**What exactly justifies the 'semi' in the presented technique? I assume it is the selection of the connected component, that represents the bone (Step II)? Please clarify.**

Indeed, the selection of connected component requires manual intervention, as does tuning of edge detection parameters. The following text has been added on Page 8 for clarification: “The semi-automated approach required manual intervention at two stages: one-time optimization of edge detection parameters for the given image contrast and resolution, and manual selection of labeled components representing the bone edges of interest. “

Results  
RE1.9:  
**You mention the duration of both, manual and semi-automatic segmentation approaches. On what kind of machine was this measured?**  
Processing times were measured using a smartphone stopwatch.

RE1.10:  
**Figure 3 and 4 use the same colors for different things (femur/tibia and semi-auto/manual) please use different colors to avoid confusion.**  
We thank the reviewer for the suggestion. The colors in Figure 4 has been modified as shown below:



RE1.11:  
**I am unsure how the data points in Fig4 are generated. From Fig3 I assume there is about 30 frames for a full flexion cycle. Fig4 only shows 10 for a half-cycle. How exactly are data from different frames (and subjects) combined?**  
Thank you for the clarification request. Regarding the frame count, Figure 3 shows 8 frames out of total 28 frames (index 0 to 27) and not 30. For Figure 4 analysis, we binned the normalized flexion cycle into 10% intervals and averaged displacement values within each bin across all datasets. This is why we have 10 data points per half cycle. The following text has been added in Section 2.3 on Page 9: ‘ For statistical analysis and visualization, the normalized percentage data were binned into 10% intervals (0-10%, 10-20%, etc), with displacement values averaged within each bin across all datasets. ‘

RE1.12:  
**Fig4: In the Horizontal displacement plots, one vertical tick/gridline represents a step of 1,25. This seems unusual and can be easily mistaken. I suggest using 1mm or 2.5mm as a ticksize.**

The minor tick size has been changed to 2.5mm as per the suggestion and can be seen in RE1.10 above.

RE1.13:  
**I would like to see the following as a supplementary material: A video of the CINE images of a full flexion at similar speed to the actual knee motion. Followed by a second full flexion, now with the segmentation overlays as in fig 3 and indication of estimated centroid positions. One subject is sufficient.**  
Please find attached the requested videos in the supplementary section

RE2.0  
**Some methodological aspects remain unclear. In addition, the range of motion in closed MRI units is relatively small (depending on the length of the lower leg). Therefore, the transferability to open MRI systems, in which the effects of bone displacement during the application of force  could also be better investigated (including orthopaedic examination techniques), should be discussed more extensively.**

RE2.1  
**The knee joint also includes the patella (or kneecap), the movement of which can be abnormal in some diseases or after injuries. Why was the patella, which is well delineated in the sagittal images, not included? Would there be more problems here?**

We thank the reviewer for this important question. We initially considered including the patella in our analysis. However, the patella exhibited less well-defined edges compared to the tibia and the femur during edge detection. Additionally, the patella had considerable through-slice motion which made it incompatible with our 2D rigid body tracking approach.

RE2.2  
**The range of motion during flexion in the knee joint in the closed MRI unit with relatively small bore (3T Siemens Prisma) is limited depending on the length of the lower leg. It should be pointed out in the paper that full-range examinations are only possible on open MRI systems, on which the proposed methods should also work.**  
We thank the reviewer for this insightful suggestion. While our current motion device would require substantial modifications for open-bore systems, the tracking algorithm itself operates directly on image data regardless of the scanner type. The algortihm’s edge detection approach could potentially maintain effectiveness even with lower signal-to-noise ratios typical of open-bore systems. We have added the following text to manuscript in the discussion section on Page 12:

“Another limitation of the current approach is the restricted range of motion inherent to closed-bore MRI systems. Our participants achieved knee flexion between 30-46 degrees, which limits assessment of deep flexion patterns. Open-bore MRI systems could provide substantially greater range of motion for more comprehensive kinematic analysis. However, such systems typically operate at lower magnetic field strengths, resulting in reduced signal-to-noise ratio and longer acquisition times. Nevertheless, the distinct contrast between cortical bone and surrounding soft tissues may provide sufficient edge definition for reliable bone boundary detection even at lower field strengths. Future validation studies in open-bore systems would confirm the robustness of our edge-based tracking approach and enable full range-of-motion analysis.”

RE2.3  
**In some orthopaedic examinations (e.g. following cruciate ligament ruptures), movements in the knee joint are measured when force is applied (e.g. anterior and posterior drawer test). Could the method also be used for quantitative assessment of such experiments?**

Although not utilized in this study, our device includes weight-loading capability, which increases the resistance during flexion-extension. This added load may accentuate pathological translation patterns in ligament-deficient knees, though this differs from the direct AP forces applied in drawer tests. Therefore, while our method could potentially detect abnormal translation patterns associated with PCL/ACL injuries, it cannot replicate or quantitatively assess drawer tests specifically, as these require controlled anterior-posterior force application rather than rotational resistance our setup provides.

RE2.4  
**How can an 'alignment error' be determined that is significantly smaller than the spatial resolution of the measurement sequences?**  
The alignment error can achieve sub-pixel precision because our optimization operates in continuous coordinate space. While the Canny edge detection produces binary edges at discrete pixel locations, our transformation parameters are continuous values. This allows the 80 reference points to be positioned at any fractional pixel coordinates. The reported alignment error (0.40 ± 0.02 mm) represents the average distance between these continuously-positioned reference points and their nearest edge pixels. We are not claiming to resolve features smaller than the voxel size, but rather optimizing alignment in continuous coordinate space.

RE2.5  
**The recording method with 2D radial GRE sequences has not become completely clear. It is described that the slice thickness is 1 mm, but the FoV is 3 mm thick. How many (sagittal) slices are acquired?**  
We thank the reviewer for noticing this mistake. It is a single 3mm thick sagittal slice acquisition, but we had mistakenly given the incorrect voxel size of [1.09x1.09x1]mm3. It has now been corrected on Page 5 to [1.09x1.09x3]mm3 with this sentence added at the end: “This acquisition protocol captured a single 3mm thick sagittal slice.”

RE2.6  
**The shape of the bones on the images changes if the lower leg has motion components perpendicular to the slice or rotational components. In this case there are no longer matching reference points on a fixed sagittal slice. How is this handled? Should layers be reconstructed from 3D data sets that depict the same sagittal plane through the lower leg at different knee flexion angles?**  
We thank the reviewer for raising this important point. Through-plane motion is indeed a limitation of our 2D tracking method. We are currently working on 3D dynamic acquisition that addresses this limitation. We have added the following text at the end of the discussion section at the end of Page 12:

“ Despite these advantages, a limitation of the current 2D approach is sensitivity to through-plane motion. While the knee motion device was designed to constrain movement to the sagittal plane, this remains a potential source of error. In cases where significant through-plane motion occurs, the bone appearances in the fixed sagittal slice change, resulting in elevated cost function values that indicate compromised tracking accuracy. Future work could address this limitation by extending the method to 3D acquisitions. “

RE2.7  
**A "semi-automated pipeline" is reported. It should be indicated at which points interventions by the examiner are necessary.**

The following text has been added in Page 8 to clarify the semi-automatic nature of the approach: “The semi-automated approach required manual intervention at two stages: one-time optimization of edge detection parameters for the given image-contrast and resolution, and manual selection of labeled components representing the bone edges of interest. “