Final project IDS702 - High Matrix

Fides Schwartz

12/12/2021

Git-Repository with all code: https://github.com/FR-Schwartz/IDS702_Final-Project

Summary

In this project, I analyze the effect of higher matrix size reconstructions (768x768 or 1024x1024 pixels) of computed tomography (CT) data compared to the current clinical standard (512x512 pixels). The hypothesis for this work is that reconstruction of image data from arterial run-off studies, that depict the arteries of the lower extremities, using high matrix sizes will improve the readability of these studies and reader confidence with either no, or only a small increase in image noise. The results of this analysis show that higher matrix reconstructions of 768x768 and 1024x1024 pixels improve image quality with higher resolution and enable more confident assessment of peripheral artery disease. The perceived image noise is not significantly increased using higher matrix reconstructions.

Introduction

Peripheral arterial disease (PAD) is defined as atherosclerotic plaque in the arteries of the lower extremity and is a common disease that affects more than 200 million people worldwide. CT angiography (CTA) is a well-established modality for imaging the peripheral vasculature. However, the diagnostic accuracy of PAD remains a challenge in certain clinical conditions.

In patients with extensive calcification of the peripheral arterial system, CT diagnostic confidence can be low, as the calcification can cause blooming artifacts, leading to the overestimation of stenosis. Similar to the coronary arteries, the arteries of the legs and feet are of small caliber so over-estimation of plaque volume plays a bigger role in mis-classification of stenosis severity than in larger arteries of the chest and abdomen.

CT scanner systems, currently in clinical use, can reconstruct data at higher matrix sizes than 512x512 at 768x768 or up to 1024x1024 pixels. It has been shown that higher matrix size reconstructions at 1024x1024 pixels can improve the image quality and assessment of lung nodules and bronchial structures over the standard reconstructions at 512x512 pixels. One drawback of the higher matrix sizes that both studies described is higher image noise but the quality of images of the lungs was rated higher for the higher matrix sizes than for the standard matrix sizes, nonetheless.

No study has compared higher matrix sizes for the imaging of peripheral artery disease to the standard clinical matrix sizes to date. In this multireader study I seek to determine whether reconstruction with a higher than standard matrix size improves image quality and clinical decision making for lower extremity CTA studies.

The questions of interest are: 1. Does the discernability of the vessel wall improve using higher matrix sizes? 2. Does perceived image noise increase using higher matrix sizes? 3. Does reader confidence in making a diagnosis based on the image improve using higher matrix sizes? 4. Is there an improvement between the intermediate matrix size of 768x768 and 1024x1024? 5. What is the inter-reader agreement between the scores for wall definition, noise levels and diagnostic confidence?

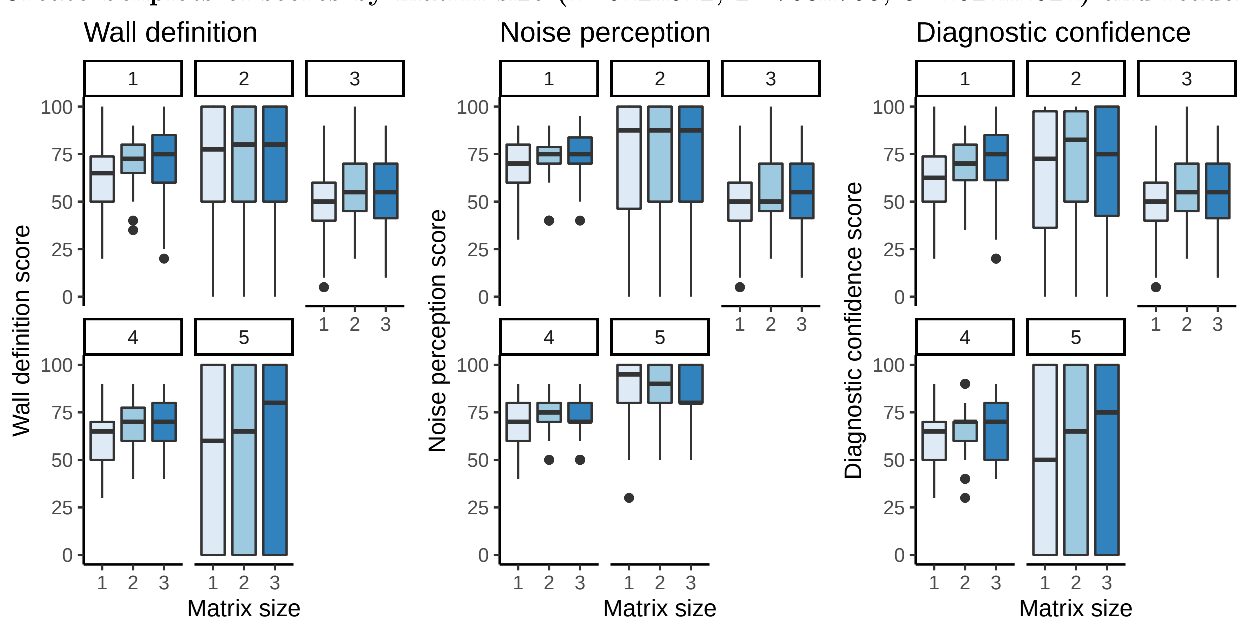
Data

The data consists of CT image data from 50 patients, that was read by five separate readers (diagnostic and interventional radiologists as well as a vascular surgeon). A representative CT slice was shown, reconstructed with the three different matrix sizes (total of 150 images), to each reader in a randomized order to avoid introducing bias. Each reader rated 1) the discernibility of the vessel wall, 2) the perceived image noise level and 3) their confidence in making a diagnosis based on the image shown on a continuous scale from 0-100. There was no missing data. This data was further sub-divided into the femoro-popliteal segments (from the groin to the bend of the knee, 15 patients) and the tibial segments (from the knee to the foot, 37 patients). This is relevant because the arteries get progressively smaller towards the feet and spatial resolution might be more relevant in the tibial segments than the femoro-popliteal ones.

Create an overall summary table of means Mean scores for the 512x512 matrix are lower than for both of the higher matrix sizes, which are very similar between themselves.

	data (N = 750)
Wall Definition	
wall_512	58.8
wall_768	64.572
wall_1024	64.5
Noise Perception	
noise_512	69.76
noise_768	72.292
noise_1024	72.18
Reader Confidence	
confidence_512	56.42
confidence_768	62.972
confidence_1024	62.6

Create boxplots of scores by matrix size (1=512x512, 2=768x768, 3=1024x1024) and reader



It seems like most of the readers gave higher mean scores for the higher matrix sizes for wall definition and diagnostic confidence but for the noise the trends are not as clear.Reader 2 is the only reader where scores did not differ clearly between matrix sizes. Reader 2 was the vascular surgeon.

Model

Model selection is performed using the overall scores for wall definition. Then the chosen model is fitted for perceived noise and diagnostic confidence for the overall data and for the sub-sets of femoro-popliteal and tibial segments.

Fitting a linear model (Model 1): This model shows statistically significant differences for both the 768x768 (P=0.02) and the 1024x1024 (P=0.02) wall definition scores compared with the standard of 512x512. The R-squared value for this model is 0.06371.

Fitting a linear mixed-effects model without slopes for matrix by reader for overall wall definition scores (Model 2): This model shows statistically significant differences for both the 768x768 (P=0.004) and the 1024x1024 (P=0.005) wall definition scores compared with the standard of 512x512. The conditional R-squared value for this model is 0.391.

Fitting a linear mixed-effects model with random slopes for matrix by reader for overall wall definition scores (Model 3): This model shows statistically significant differences for both the 768x768 (P=0.005) and the 1024x1024 (P=0.005) wall definition scores compared with the standard of 512x512.

Generate QQ-plots for wall definition for all models The assumption of normality holds for all three models with most of the residuals distributed along the 45-degree line of the Q-Q-plot, though there is more deviance from the 45-degree line in Model 1 than in both Model 2 and 3 (see appendix). Since Model 2 and 3 are performing better and are linear mixed effects models, the assumptions of linearity and equal variance are not tested. The assumption of independence cannot be tested but based on how the data was acquired this assumption holds.

Compare hierarchical vs linear mixed effects model AIC is lower for the linear mixed effects model without (6931.4) and with random slopes (6943.2) for matrix by reader than for the standard linear model (7118.8). In addition to the AIC, the BIC is lower for the model without random slopes (6959.2) than for the one with random slopes (6998.6), even though the chi-squared test shows no significant difference. Thus, random slopes will not be added to the final model (model 2).

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
wall_all2	6.00	6931.44	6959.16	-3459.72	6919.44			
wall_all3	12.00	6943.18	6998.62	-3459.59	6919.18	0.26	6	0.9997

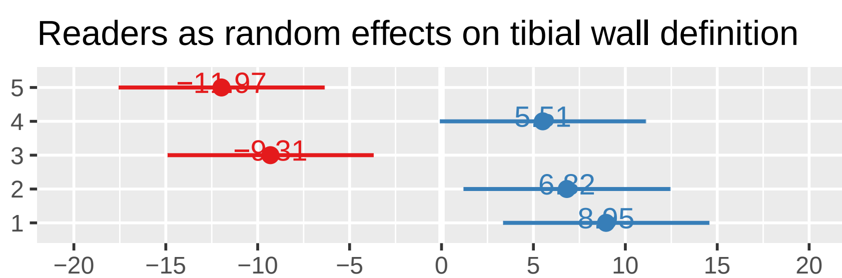
The simple linear mixed-effects model without random slopes uses a regression for the scores of each wall definition, noise level and reader confidence as outcome and subject/patient, reader and matrix as variables. Scores are compared between matrix reconstructions (overall and separately for tibial and femoro-popliteal arteries) with matrix treated as a fixed effect and subject and reader as random effects. This model assumes that the residual error has a Gaussian distribution:

$$Model2 < -lmer(wall\ matrix + (1|subject) + (1|reader) + (matrix - 1|reader), data = data)$$

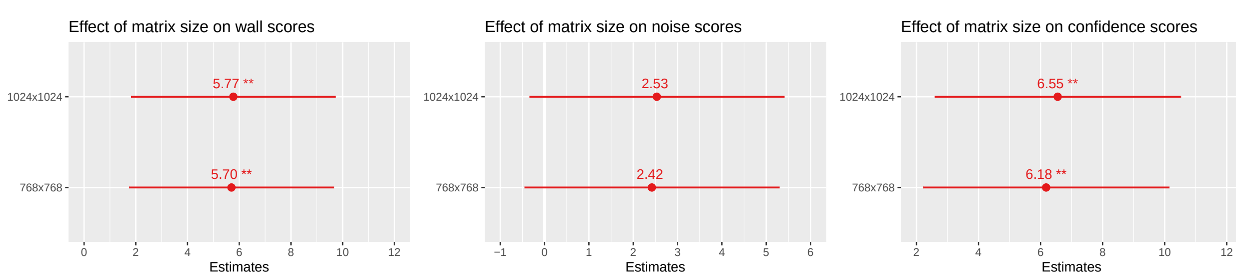
Fit linear mixed effects models without slopes for scores overall: The assumption of normality holds for both noise perception and diagnostic confidence, though there is more deviation from the 45-degree line for noise perception than the other two scores (see appendix). The predicted values for all three parameters are similar to the true values, with higher mean scores for the higher matrix sizes of wall definition and diagnostic confidence, though there is a clearer trend for higher scores for noise in the predictions that was not seen as clearly when plotting the individual reader responses. Predicted values (marginal effects) for wall definition and diagnostic confidence are similar to the observed reader scores. The predicted values of noise perception do seem to show higher scores with the higher matrix sizes, but this is not statistically significant (see appendix). The variance in scores for overall wall definition and diagnostic confidence is not explained by the random effects of reader (11% and 8%) or subject (34% and 36%) but the fixed effects of higher matrix reconstructions reach significance (768: 0.004 and 0.001, 1024: 0.005 and 0.002). The variance in scores for overall perceived noise is not explained by reader (34%) or subject (32%) and the matrix sizes do not reach significance.

Fit linear mixed effects models without slopes only for tibial arteries: Using the Model 2 structure for just the segments classified as "tibials". The assumption of normality holds for all three parameters (see appendix). The deviation from the 45-degree line is decreased compared to the overall scores. The variance in scores for tibial wall definition and diagnostic confidence is not explained by the random effects of reader (15% and 13%) or subject (20% and 22%) but the fixed effects of higher matrix reconstructions reach significance (768: 0.002 and 0.0003, 1024: 0.02 and 0.006). The variance in scores for tibial perceived noise is not explained by reader (40%) or subject (26%) and the matrix sizes do not reach significance.

Fit linear mixed effects models without slopes for femoro-popliteal arteries: Using the Model 2 structure for just the segments classified as "fempop". The assumption of normality holds for all three parameters (see appendix). The deviation from the 45-degree line is slightly increased compared to the overall scores. The random effects of reader (16%) and subject (35%) do not explain the variation in femoropopliteal wall definition scores but the matrix size of 1024x1024 reaches the level of significance (0.04). The score variation in perceived noise and diagnostic confidence is not explained by reader (42% and 11%) or subjects (26% and 42%), and the matrix sizes do not reach significance either. The random effects plots for readers are included in the appendix and an exemplary plot of wall definition in the tibial arteries is shown:



In addition the effects of matrix size on overall wall definition scores are shown (plots for tibial and femoropopliteal segments in appendix):



The effect of matrix size on all three parameters is shown in table format in the appendix.

Overall ICC scores for wall definition (0.4) and diagnostic confidence (0.4) indicate poor correlation while the scores for perceived noise indicate moderate correlation (0.5). A similar pattern is seen for the tibial segments (wall ICC=0.3, noise ICC=0.5, confidence ICC=0.3). The correlations are slightly better for the femoropopliteal segments (wall ICC=0.4, noise ICC=0.5, confidence ICC=0.5).

Conclusions

1. Does the discernability of the vessel wall improve using higher matrix sizes? Reconstructions with a 768x768 and a 1024x1024 matrix received a higher mean score for wall definition than a 512x512 matrix. The difference in scores between each of the higher matrix sizes and the standard matrix size was statistically significant at P=0.004 and P=0.005. This effect is more pronounced for the tibial segments and not as clear for the femoropopliteal segments.

2. Does perceived image noise increase using higher matrix sizes? The perceived image noise was not significantly different with a 768x768 and a 1024x1024 matrix size compared to a 512x512 matrix size. The level of significance was not reached for either matrix size. This is similar for the tibial and femoropopliteal segments.

3. Does reader confidence in making a diagnosis based on the image improve using higher matrix sizes? Reader confidence in determining grade of stenosis was significantly higher for 768x768 and 1024x1024 matrix sizes compared to a 512x512 matrix size. This effect is more pronounced for the tibial segments and not as clear for the femoropopliteal segments.

4. Is there an improvement between the intermediate matrix size of 768x768 and 1024x1024? There is no significant difference between wall definition scores or diagnostic confidence scores between the two higher matrix sizes.

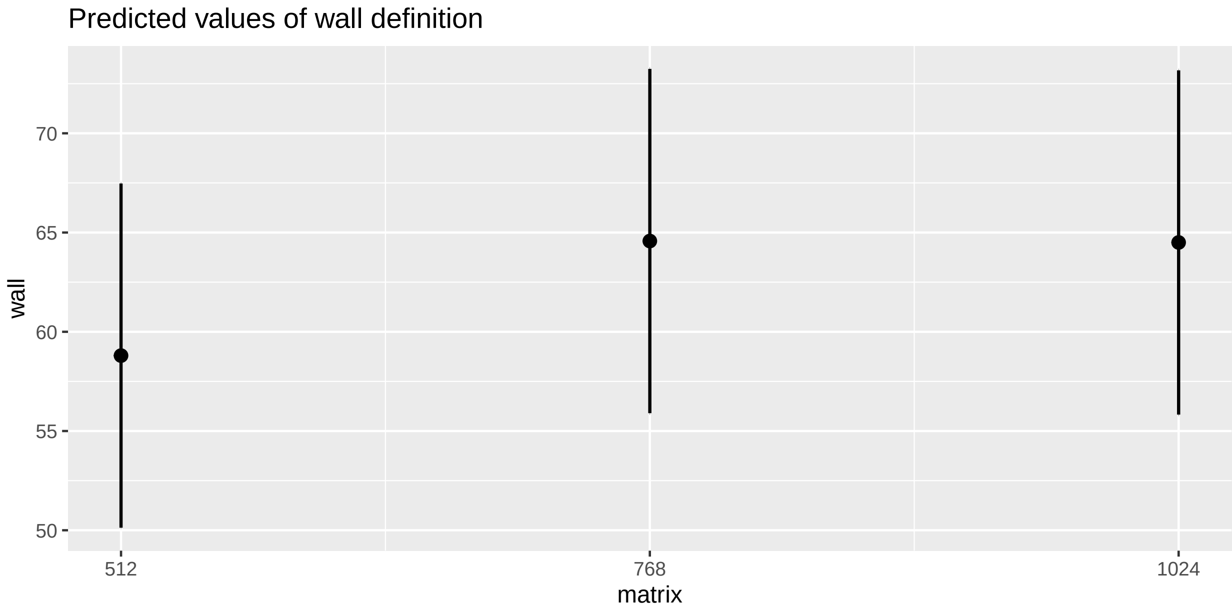
5. What is the inter-reader agreement between the scores for wall definition, noise levels and diagnostic confidence? The rho is 0.5 for inter-reader agreement in all of the scores, this translates to a kappa of 0.5, which is generally interpreted as moderate inter-reader agreement. The ICC scores are not as convincing with poor agreement for most parameters and moderate agreement for noise levels overall and in the sub-segments, and moderate agreement for diagnostic confidence in the femoral segments.

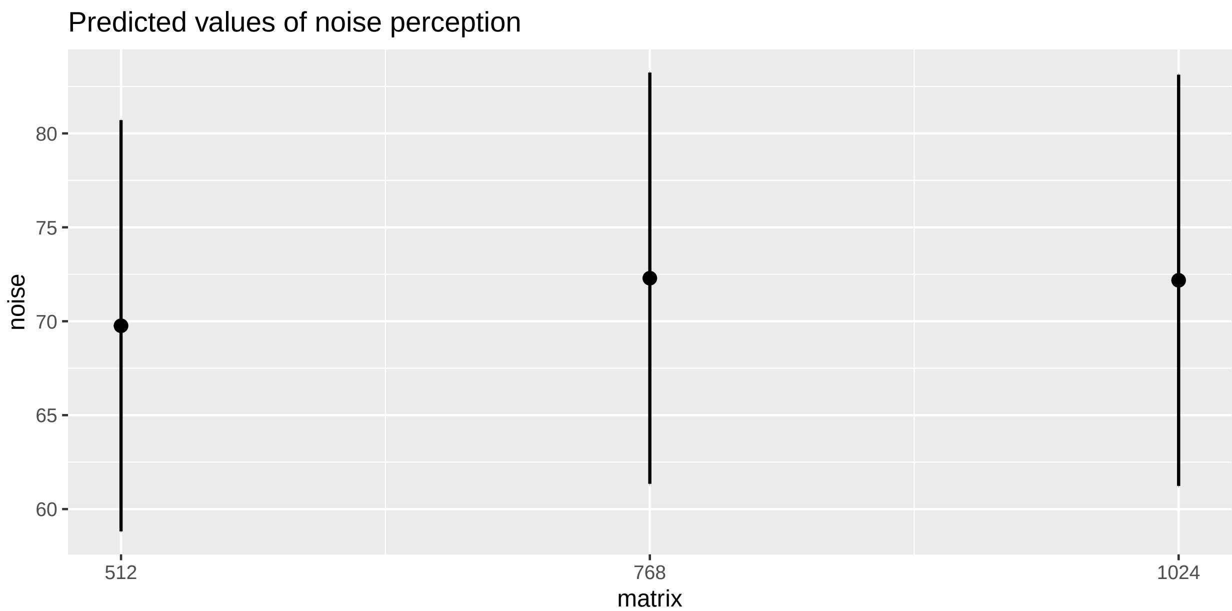
In conclusion, higher matrix reconstructions of 768x768 and 1024x1024 pixels improve image quality over standard reconstructions at 512x512 pixels with higher resolution and enable more confident assessment of peripheral artery disease, based on the analysis of this data. This is more pronounced in the smaller arteries below the knee than in the larger arteries above the knee, so higher matrix reconstructions might only be necessary for the lower part of the legs. No significant differences were observed between the two higher matrix reconstructions, so an intermediate matrix size of 768x768 could be used in clinical practice to save on reconstruction time and data storage requirements.

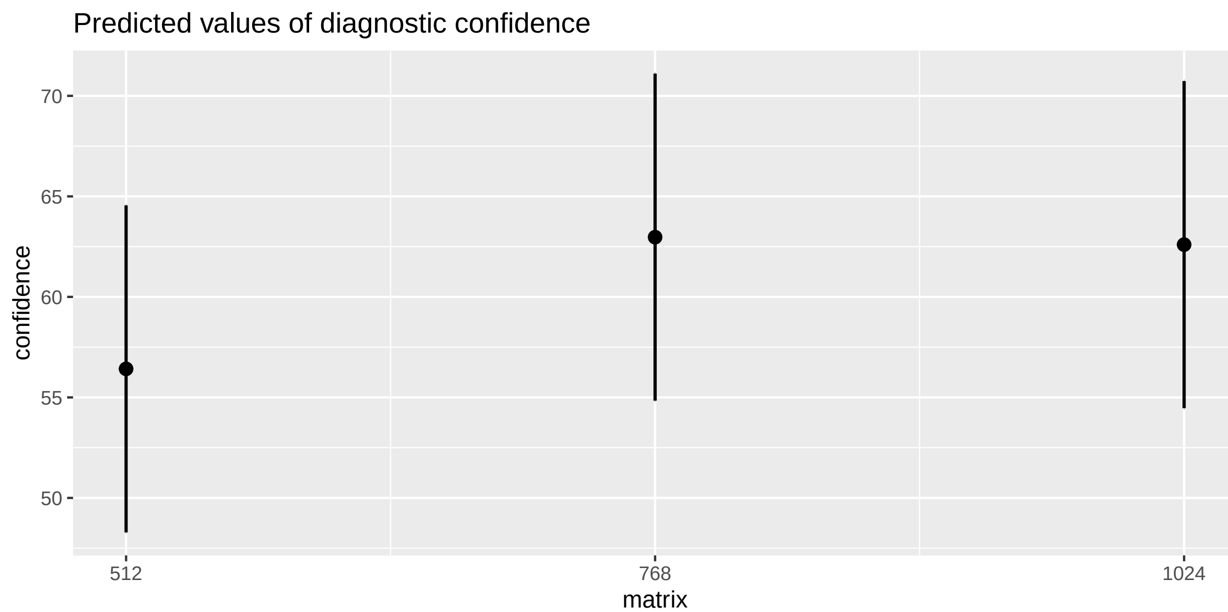
Limitations The ICC scores for all models are relatively low and indicate poor correlation based on the guidelines given by Koo and Li (2016). This is in agreement with the exploratory data analysis box plots that showed one reader (reader 2=vascular surgeon) did not give very different scores for any of the parameters between matrix sizes. This might indicate that radiologists who are more used to working with CT data profit more from higher matrix size reconstructions than vascular surgeons, who are used to working with angiographic data (that does not depict the vessel wall). This would have to be explored with a larger population of readers, who are vascular surgeons, or alternatively, the CT data could be reconstructed in post-processing to resemble an angiogram more closely.

Appendix

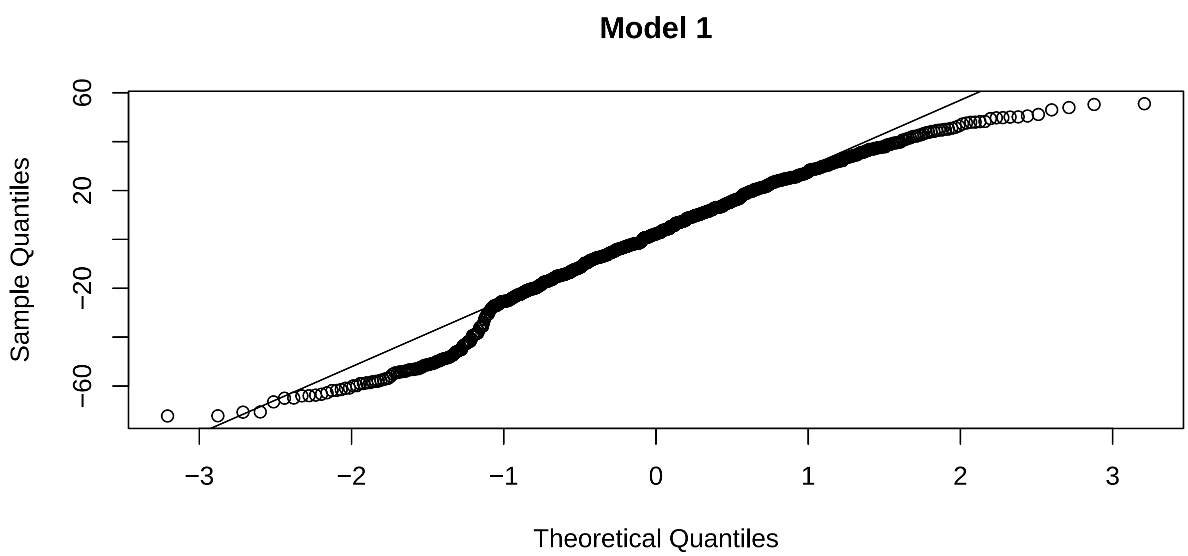
Predicted values for wall definition, noise perception and diagnostic confidence



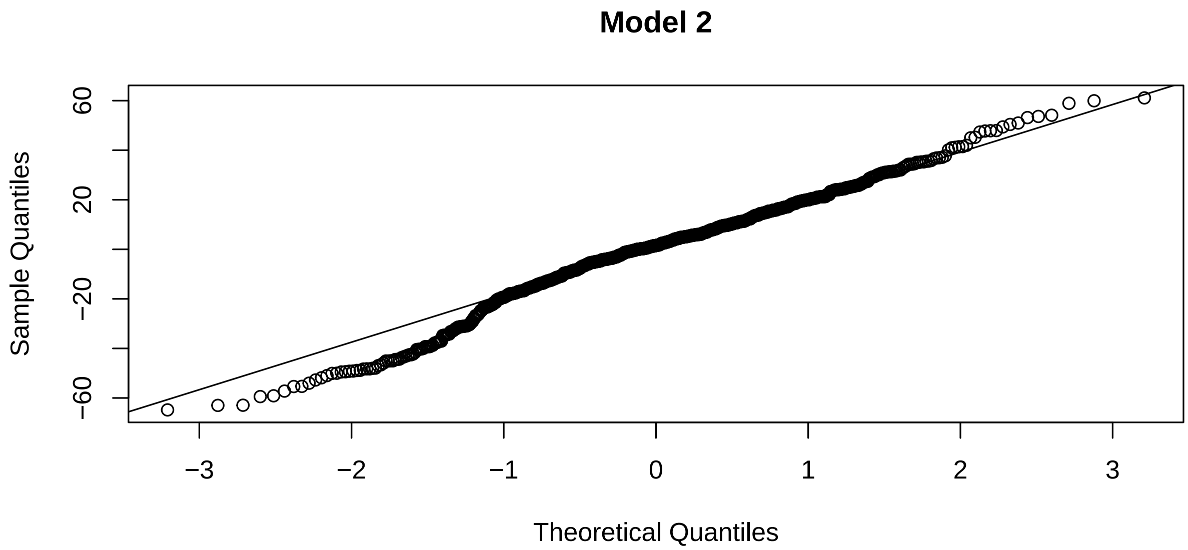




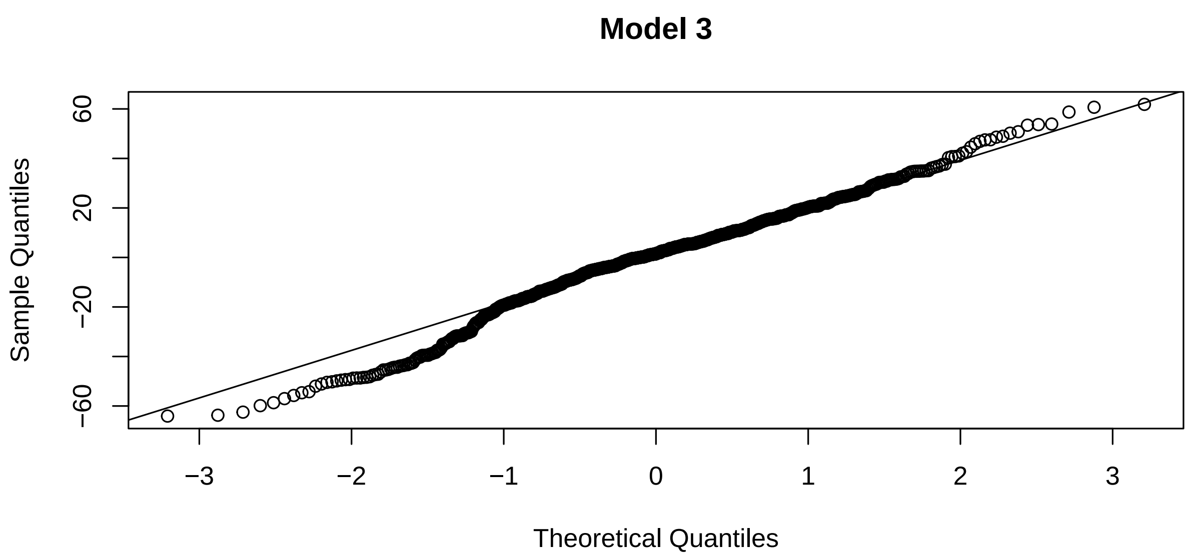
QQ-plots QQ-plot for overall wall definition scores model 1 (linear model)



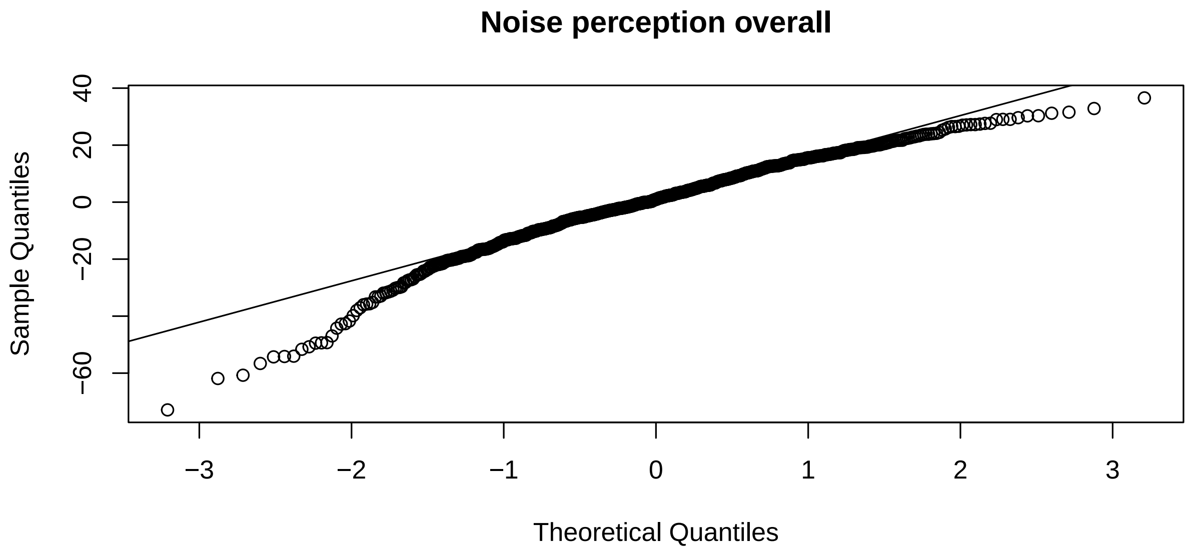
QQ-plot for overall wall definition scores model 2 (linear mixed effects model)



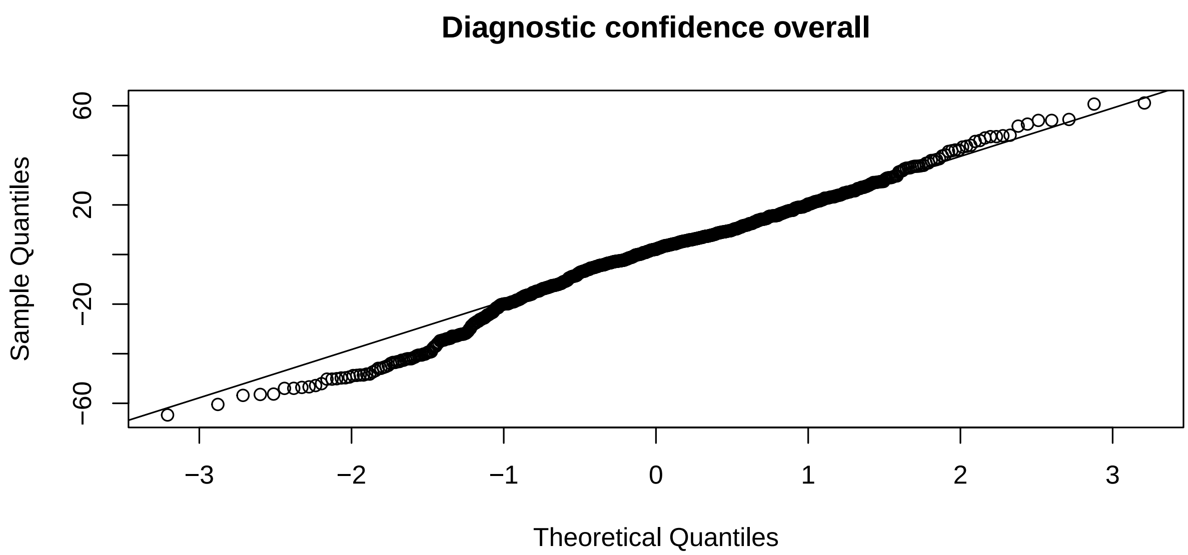
QQ-plot for overall wall definition scores model 3 (linear mixed effects model with varying slopes)



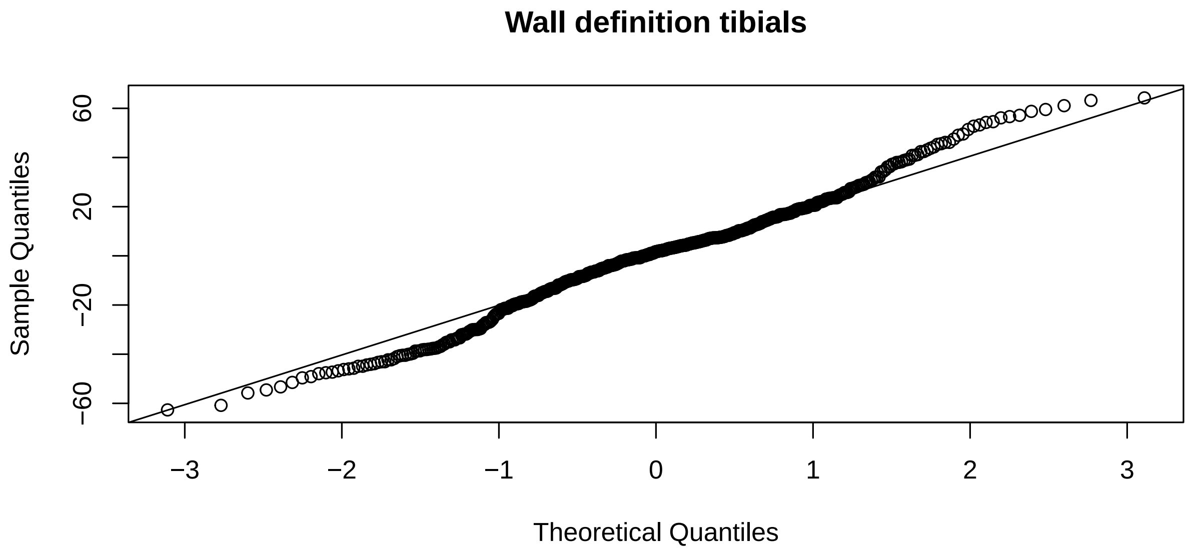
QQ-plot for model 2 overall noise perception



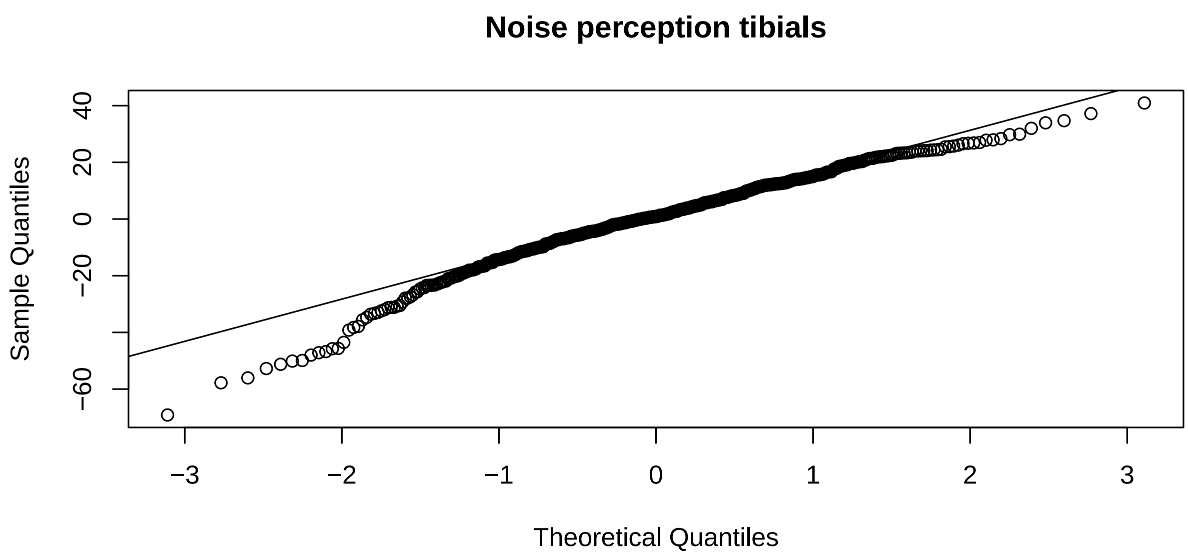
QQ-plot for model 2 overall diagnostic confidence



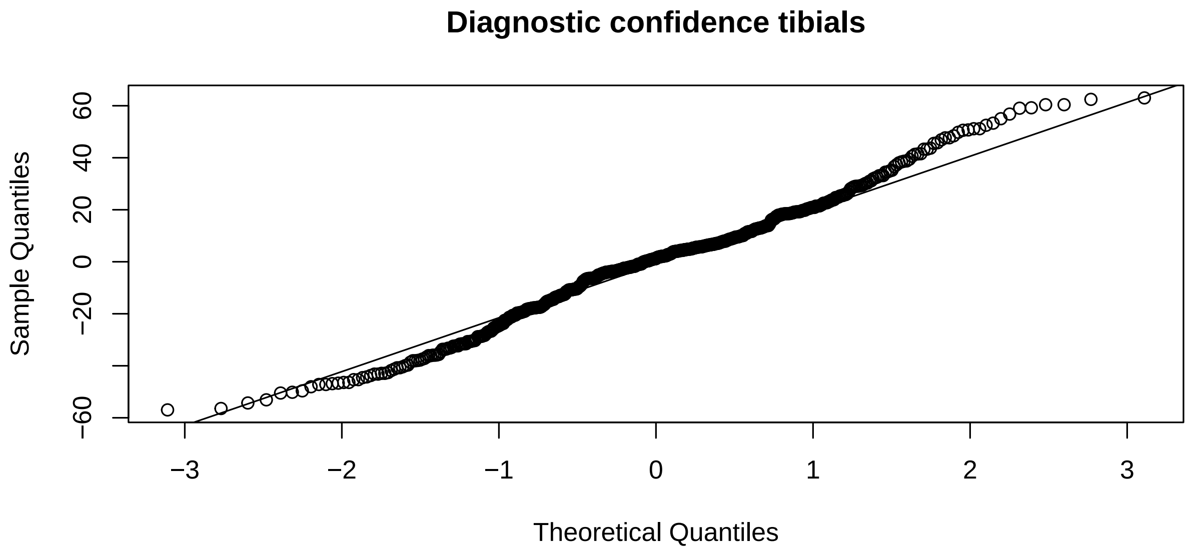
QQ-plot for model 2 wall definition in tibial segments



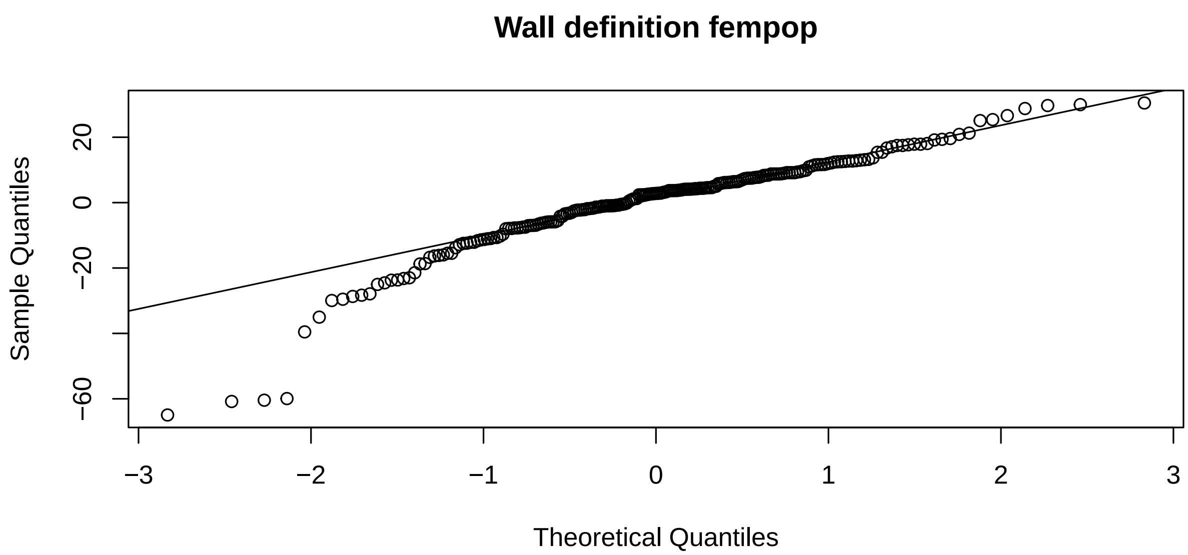
QQ-plot for model 2 noise perception in tibial segments



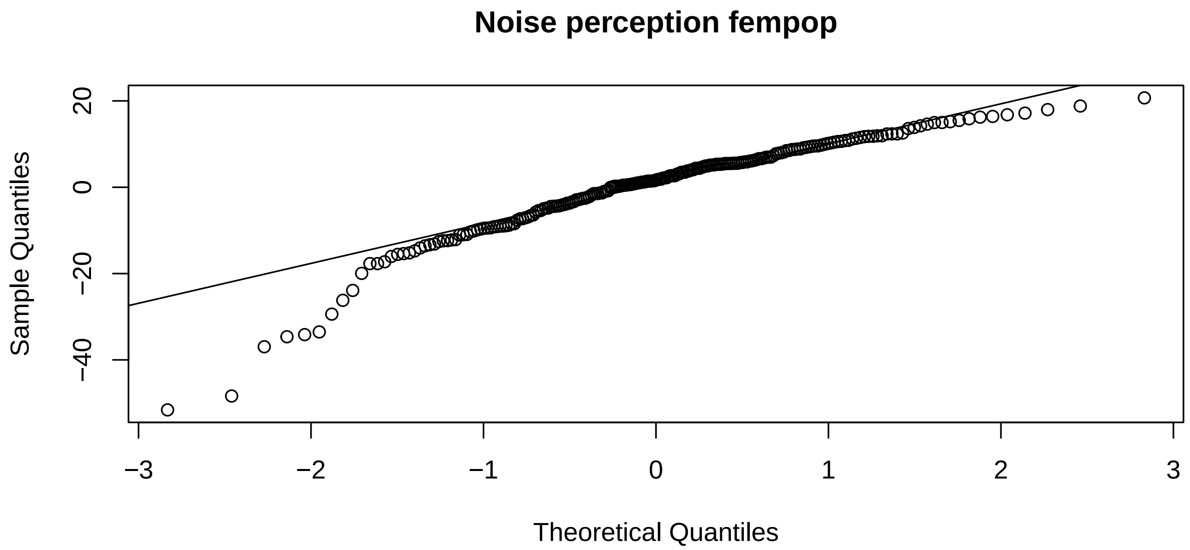
QQ-plot for model 2 diagnostic confidence in tibial segments



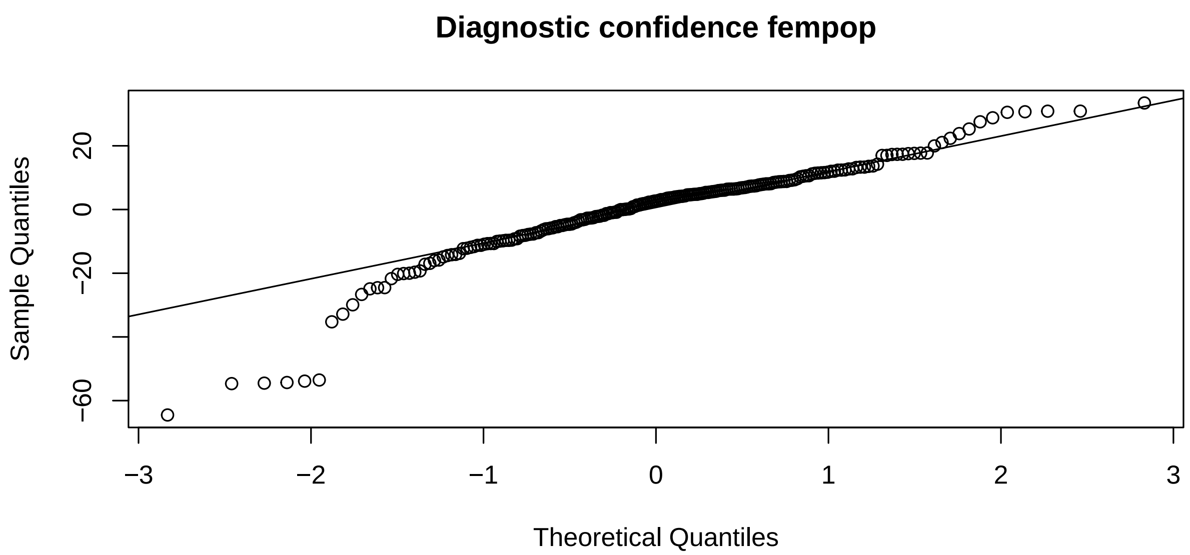
QQ-plot for model 2 wall definition in femoropopliteal segments



QQ-plot for model 2 noise perception in femoropopliteal segments

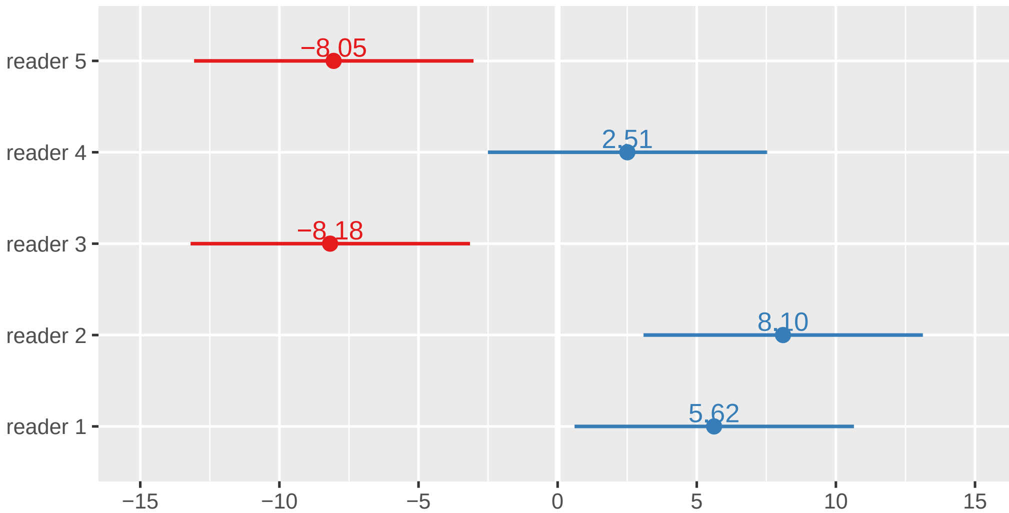


QQ-plot for model 2 diagnostic confidence in femoropopliteal segments

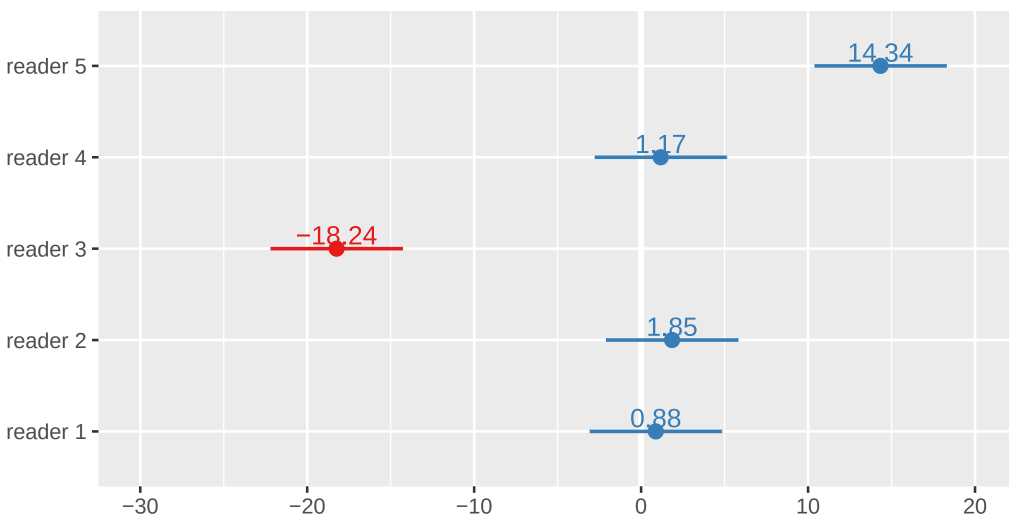


Readers as random effects on overall scores and on scores separated out for tibial and femoropopliteal segments in plots

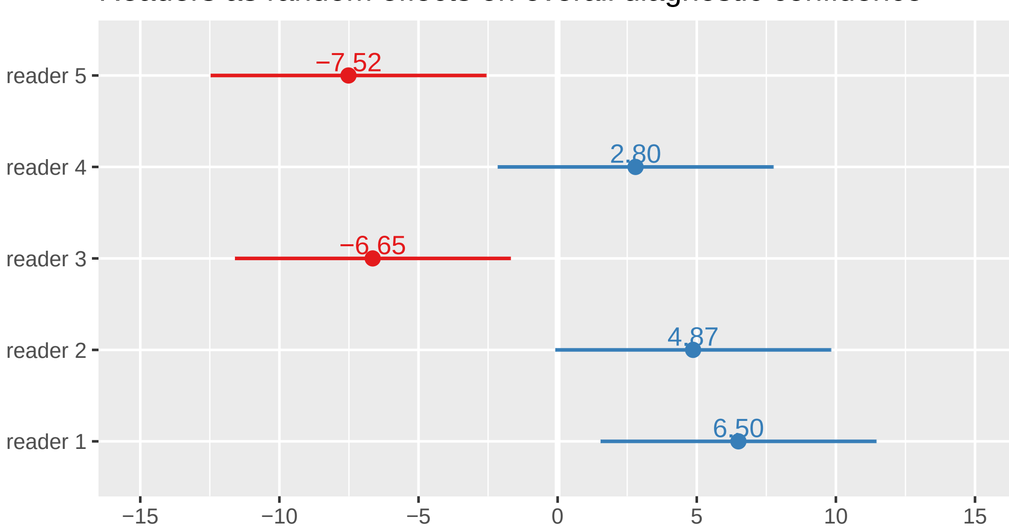
Readers as random effects on overall wall definition



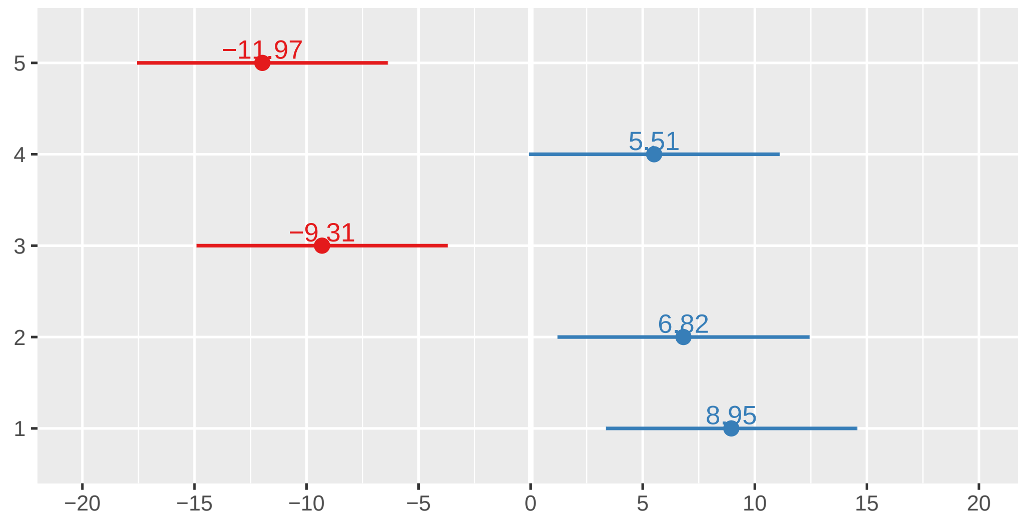
Readers as random effects on overall noise perception



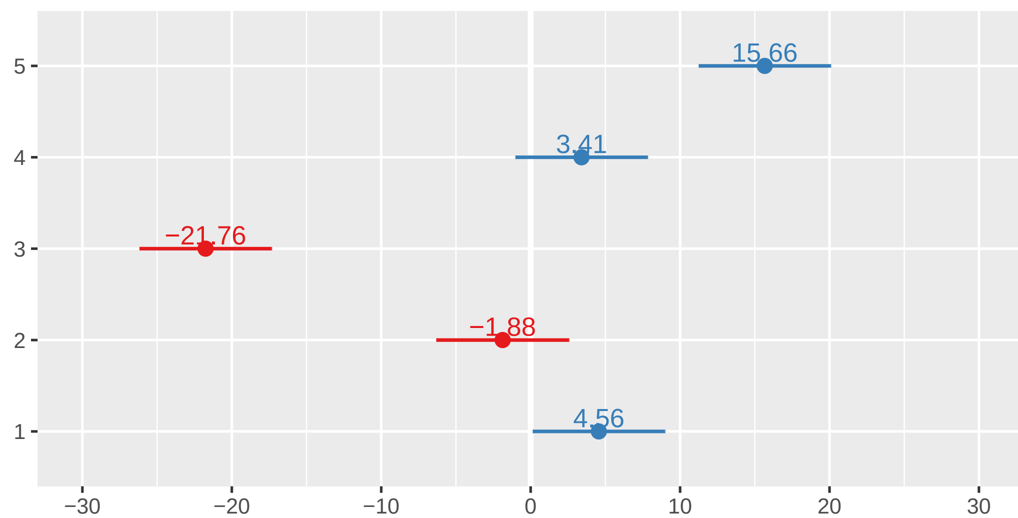
Readers as random effects on overall diagnostic confidence



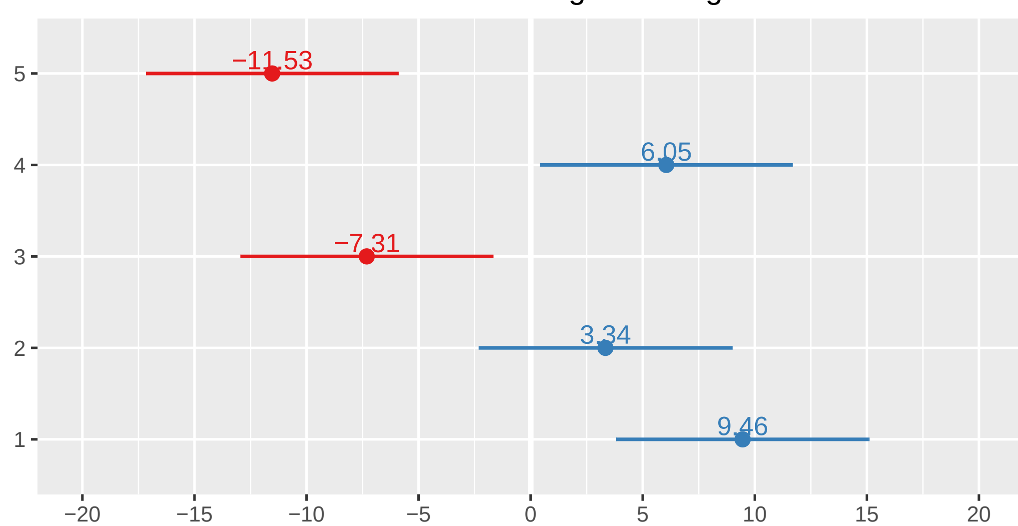
Readers as random effects on tibial wall definition

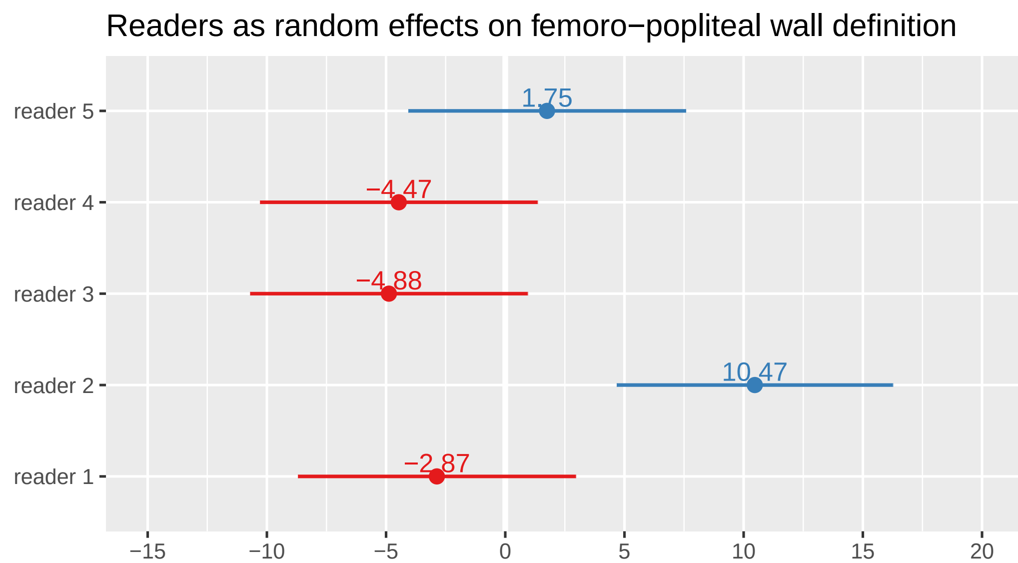


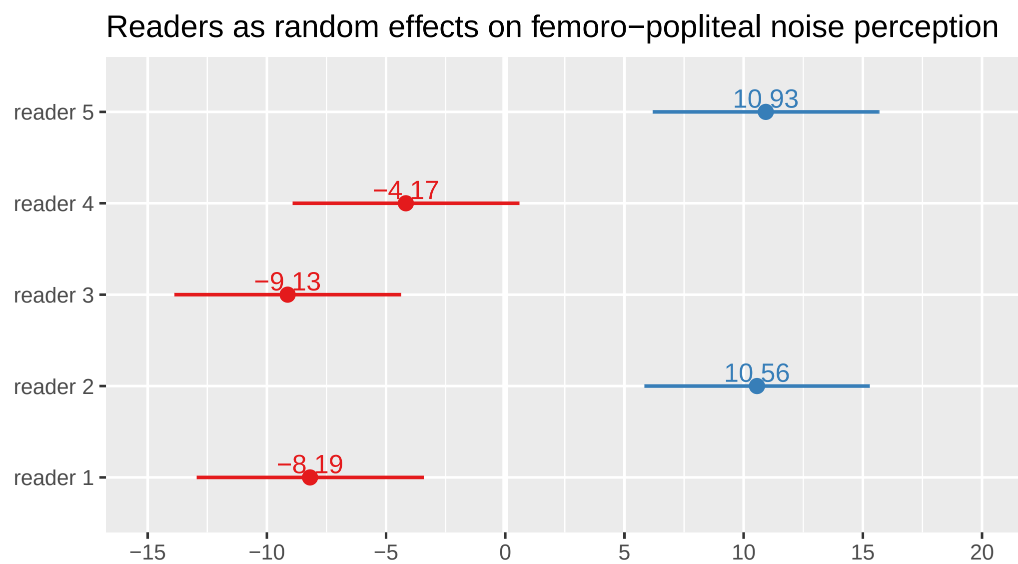
Readers as random effects on tibial segment noise perception

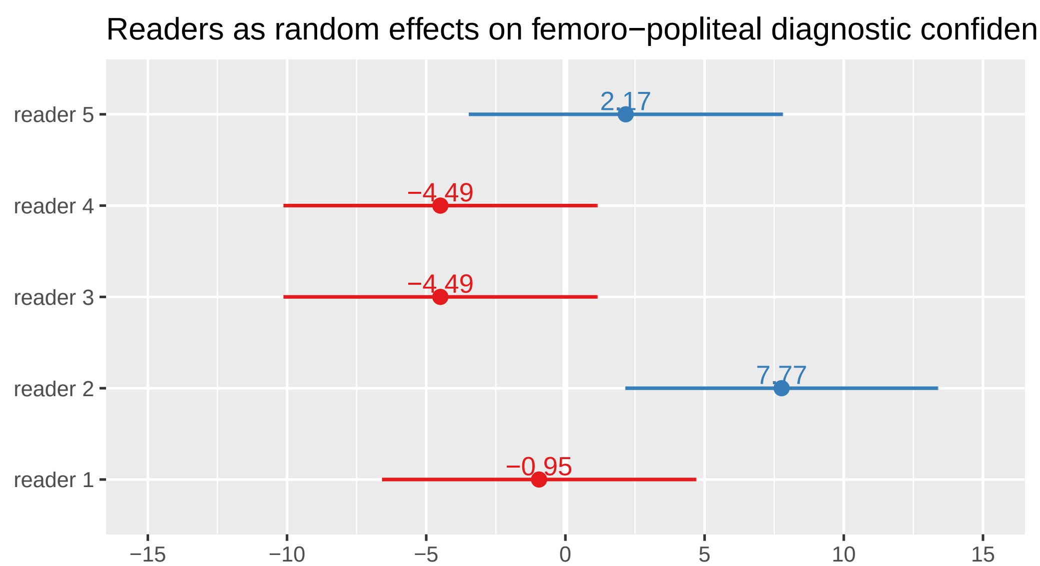


Readers as random effects on tibial segment diagnostic confidence




Readers as random effects on femoro−popliteal wall definition
reader 5
reader 4
reader 3
reader 2
reader 1
1.75
−4.47
−4.88
10.47
−2.87
−15
−10
−5
0
5
10
15
20


Readers as random effects on femoro−popliteal noise perception
reader 5
reader 4
reader 3
reader 2
reader 1
10.93
−4.17
−9.13
10.56
−8.19
−15
−10
−5
0
5
10
15
20


Readers as random effects on femoro−popliteal diagnostic confiden
reader 5
reader 4
reader 3
reader 2
reader 1
2.17
−4.49
−4.49
7.77
−0.95
−15
−10
−5
0
5
10
15

Effect of matrix size on tibial and femoropopliteal scores in plots

Tables of effect of matrix size on overall, tibial and femoropopliteal segment scores

	Effects of matrix size on overall wall definition scores		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	58.80	50.13 – 67.47	**<0.001**
768x768	5.77	1.82 – 9.73	**0.004**
1024x1024	5.70	1.75 – 9.65	**0.005**
Random Effects			
σ^2	508.51		
$\tau_{00\ subject}$	257.13		
$\tau_{00\ reader}$	62.00		
ICC	0.39		
$N_{subject}$	50		
N_{reader}	5		
Observations	750		
Marginal R^2 / Conditional R^2	0.009 / 0.391		

	Effects of matrix size on tibial wall definition scores		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	50.88	40.59 – 61.16	**<0.001**
768x768	8.12	3.10 – 13.14	**0.002**
1024x1024	6.27	1.25 – 11.29	**0.014**
Random Effects			
σ^2	581.06		
$\tau_{00\ subject}$	143.40		
$\tau_{00\ reader}$	101.90		
ICC	0.30		
$N_{subject}$	37		
N_{reader}	5		
Observations	534		
Marginal R^2 / Conditional R^2	0.014 / 0.307		

	Effects of matrix size on femoro-poplietal wall definition scores		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	77.65	68.54 – 86.77	**<0.001**
768x768	0.39	-4.72 – 5.50	0.881
1024x1024	5.48	0.31 – 10.65	**0.038**
Random Effects			
σ^2	241.69		
$\tau_{00\ subject}$	132.75		
$\tau_{00\ reader}$	46.16		
ICC	0.43		
$N_{subject}$	15		
N_{reader}	5		
Observations	216		
Marginal R^2 / Conditional R^2	0.015 / 0.434		

Predictors	Effects of matrix size on overall noise perception scores Estimates	CI	p
(Intercept)	69.76	58.81 – 80.71	**<0.001**
768x768	2.53	-0.34 – 5.40	0.084
1024x1024	2.42	-0.45 – 5.29	0.099
Random Effects			
σ^2	268.30		
$\tau_{00\ subject}$	128.87		
$\tau_{00\ reader}$	137.73		
ICC	0.50		
$N_{subject}$	50		
N_{reader}	5		
Observations	750		
Marginal R^2 / Conditional R^2	0.003 / 0.500		

Predictors	Effects of matrix size on tibial noise perception scores Estimates	CI	p
(Intercept)	65.43	52.64 – 78.21	**<0.001**
768x768	3.25	-0.27 – 6.77	0.070
1024x1024	2.12	-1.40 – 5.64	0.239
Random Effects			
σ^2	285.76		
$\tau_{00\ subject}$	98.78		
$\tau_{00\ reader}$	191.31		
ICC	0.50		
$N_{subject}$	37		
N_{reader}	5		
Observations	534		
Marginal R^2 / Conditional R^2	0.003 / 0.505		

Predictors	Effects of matrix size on femoro-poplietal noise perception scores Estimates	CI	p
(Intercept)	80.54	70.56 – 90.52	**<0.001**
768x768	0.38	-3.52 – 4.28	0.847
1024x1024	3.19	-0.75 – 7.14	0.113
Random Effects			
σ^2	141.03		
$\tau_{00\ subject}$	49.44		
$\tau_{00\ reader}$	102.86		
ICC	0.52		
$N_{subject}$	15		
N_{reader}	5		
Observations	216		
Marginal R^2 / Conditional R^2	0.007 / 0.522		

Predictors	Effects of matrix size on overall diagnostic confidence scores Estimates	CI	p
(Intercept)	56.42	48.28 – 64.56	**<0.001**
768x768	6.55	2.59 – 10.51	**0.001**
1024x1024	6.18	2.22 – 10.14	**0.002**
Random Effects			
σ^2	510.64		
$\tau_{00\ subject}$	291.47		
$\tau_{00\ reader}$	46.78		
ICC	0.40		
$N_{subject}$	50		
N_{reader}	5		
Observations	750		
Marginal R^2 / Conditional R^2	0.011 / 0.405		

	Effects of matrix size on tibial diagnostic confidence scores		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	47.96	38.19 – 57.74	**<0.001**
768x768	9.24	4.23 – 14.25	**<0.001**
1024x1024	7.11	2.10 – 12.12	**0.005**
Random Effects			
σ^2	579.08		
$\tau_{00\ subject}$	162.82		
$\tau_{00\ reader}$	86.03		
ICC	0.30		
$N_{subject}$	37		
N_{reader}	5		
Observations	534		
Marginal R^2 / Conditional R^2	0.019 / 0.314		

	Effects of matrix size on femoro-poplietal diagnostic confidence scores		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	80.54	70.56 – 90.52	**<0.001**
768x768	0.38	-3.52 – 4.28	0.847
1024x1024	3.19	-0.75 – 7.14	0.113
Random Effects			
σ^2	141.03		
$\tau_{00\ subject}$	49.44		
$\tau_{00\ reader}$	102.86		
ICC	0.52		
$N_{subject}$	15		
N_{reader}	5		
Observations	216		
Marginal R^2 / Conditional R^2	0.007 / 0.522		