Forensic Analysis of Insole Footprints Reliability of a New Method

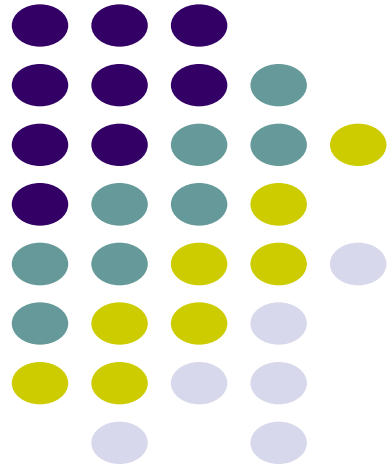
Dr Nick Howsam

PhD, MSc, BSc (Hons), BSc (Hons), PG Dip, Dip FHID

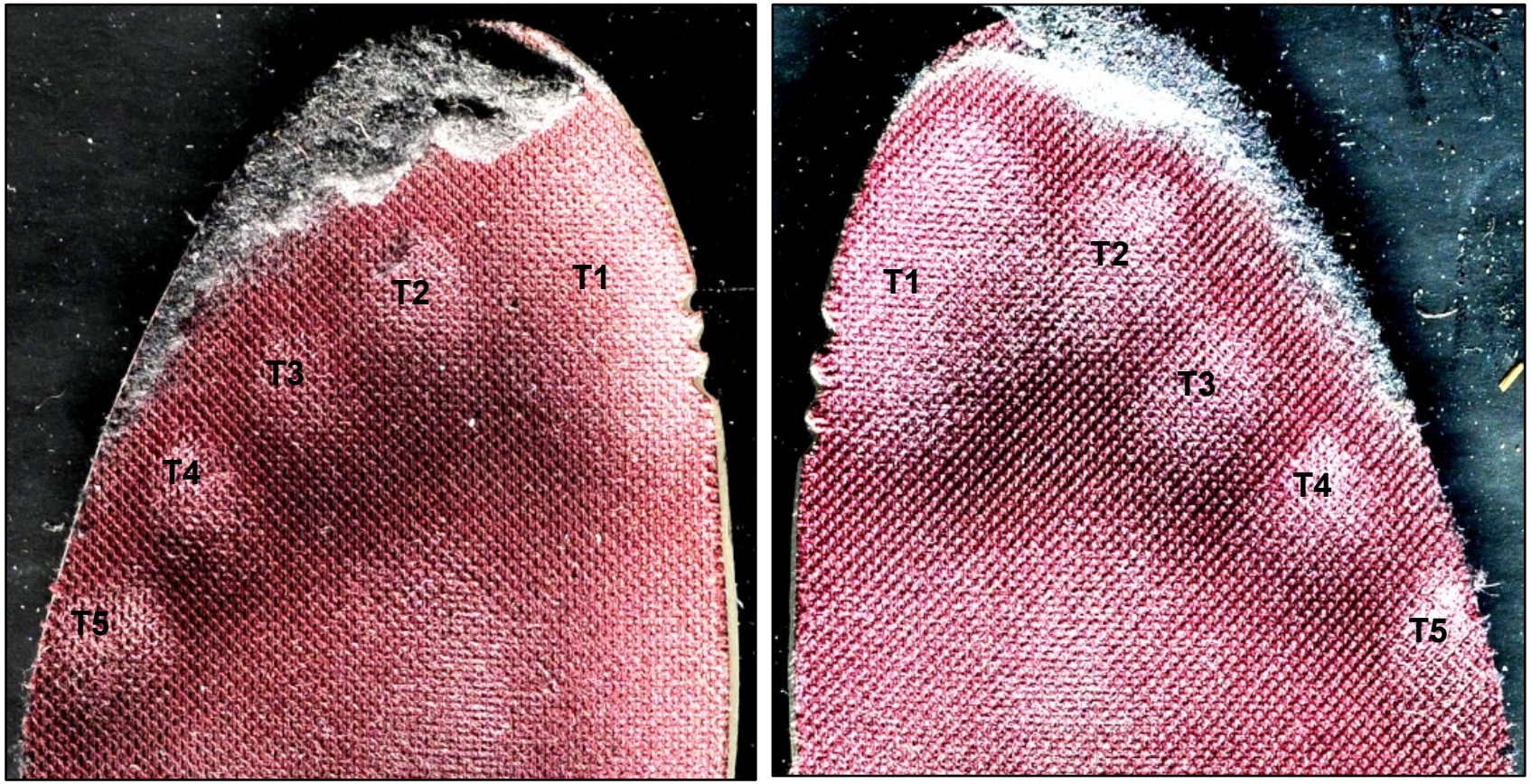
Podiatrist

Forensic Podologist

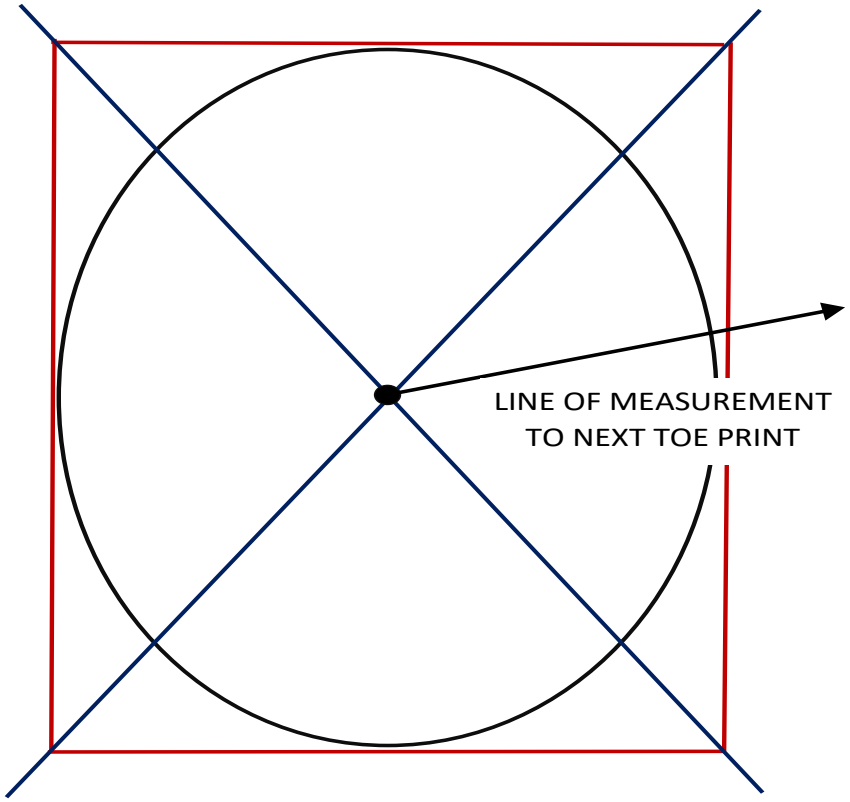
Expert Witness – Defence (Civil Courts)

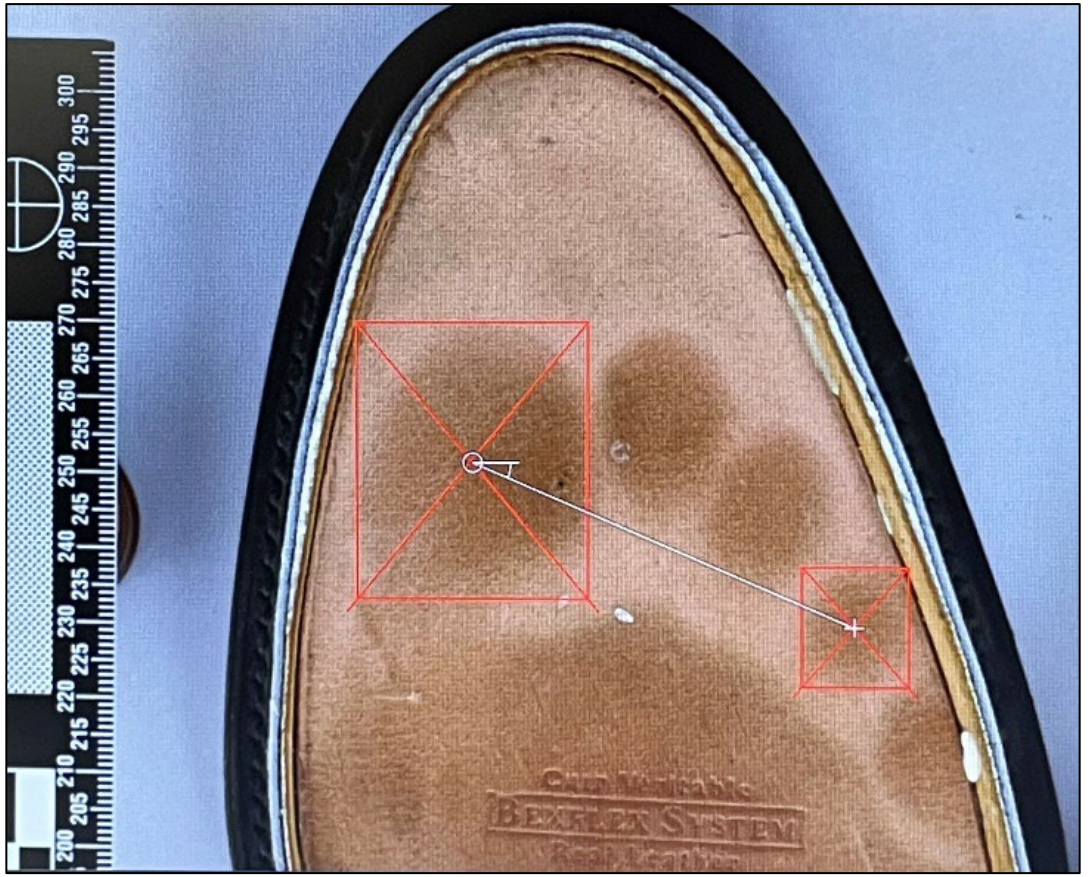


Howsam-Frame (HF) Method



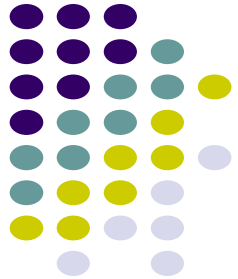
Howsam-Frame (HF) Method





GNU Image Manipulation Programme (GIMP)

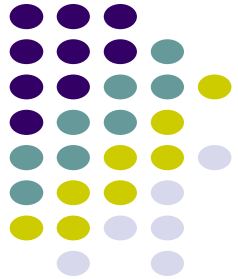
Forensic Science Regulator (FSR)



- Forensic methods used in evidential casework and submitted to the Criminal Justice System, must be validated prior to their use (Forensic Science Regulator, 2020)

- Must show that methods are fit for purpose, that is, conclusions generated from their use can be relied upon in courts of law (Forensic Science Regulator, 2020)

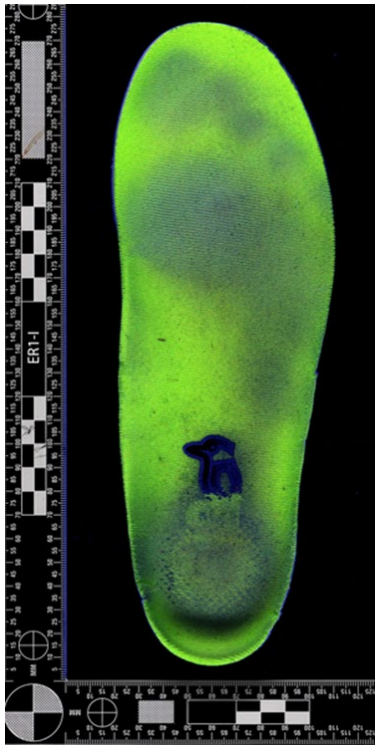
Reliability for Forensic Practice



- Is the method valid for forensic practice?
- Will the method produce reliable results?

Evaluation of Reliability

- What is reliability?
- How do I demonstrate reliability?

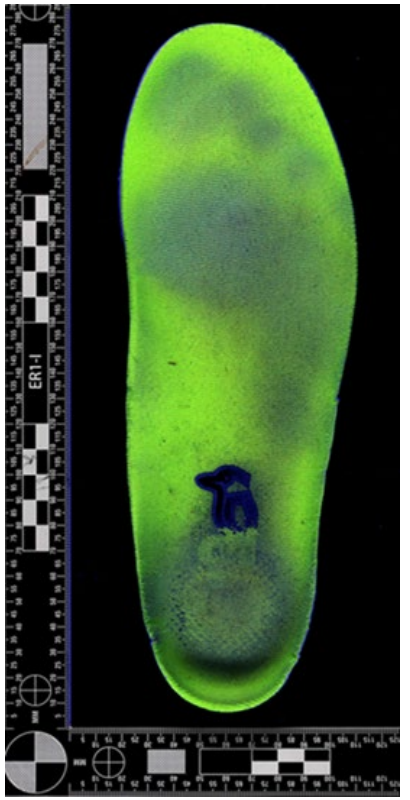


What is reliability?

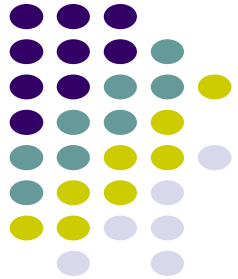
- How consistent a method is in generating a result

- Reliability:

  True Value + Measurement Error

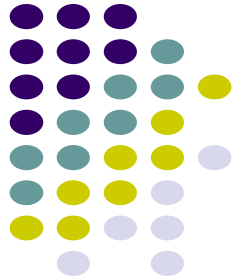


How do I demonstrate reliability?



- Intra-rater Reliability
- **Repeatability**: need to know I can repeat measurements consistently
- Inter-rater Reliability
- **Reproducibility**: need to know more than one practitioner can reproduce measurements consistently

How do I demonstrate reliability?



- Statistical Tests – **Triangulate**
- Intra Class Correlation Coefficient (ICC)
- Standard Error of Measurement (SEM)
- Bland & Altman's 95% Limits of Agreement (LOA)

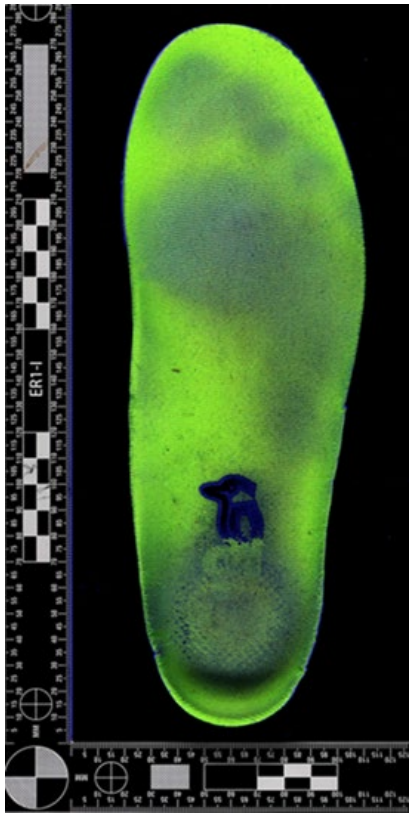
Interpretation

Intra Class Correlation Coefficient (ICC)

<0.5 indicative of poor reliability
0.5 to 0.75 indicative of moderate reliability
0.75 to 0.9 indicative of good reliability
>0.90 indicative of excellent reliability

According to Portney and Watkins (2000) & Koo and Li (2016)

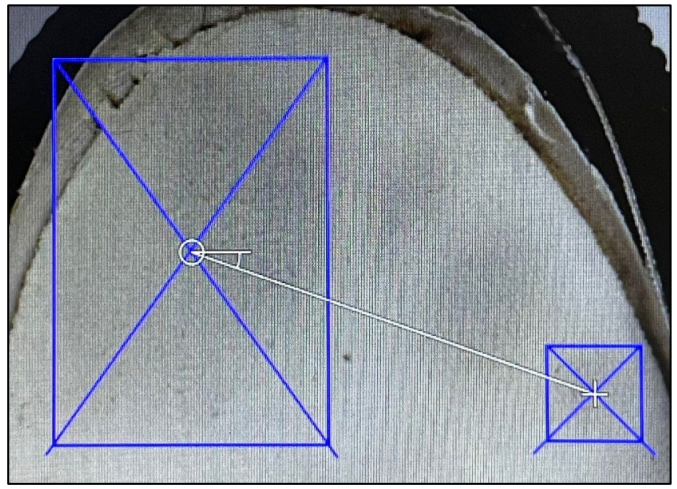
About PRECISION

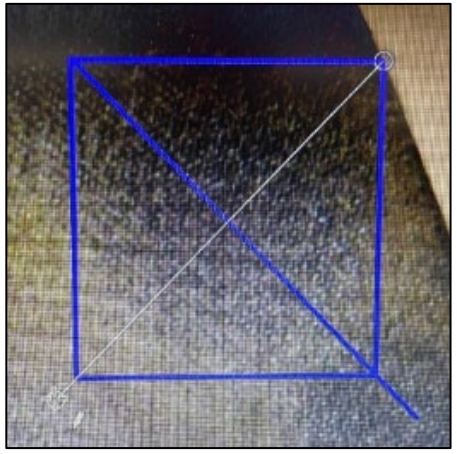


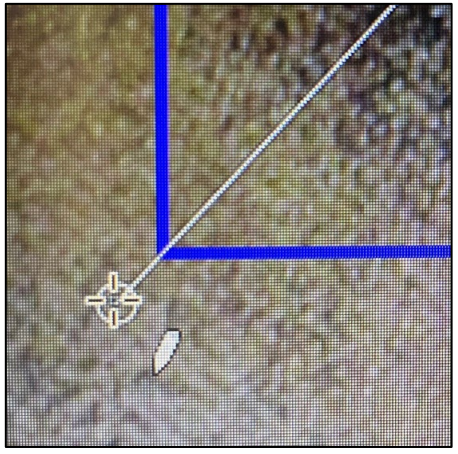
Not ACCURACY

Pilot Study

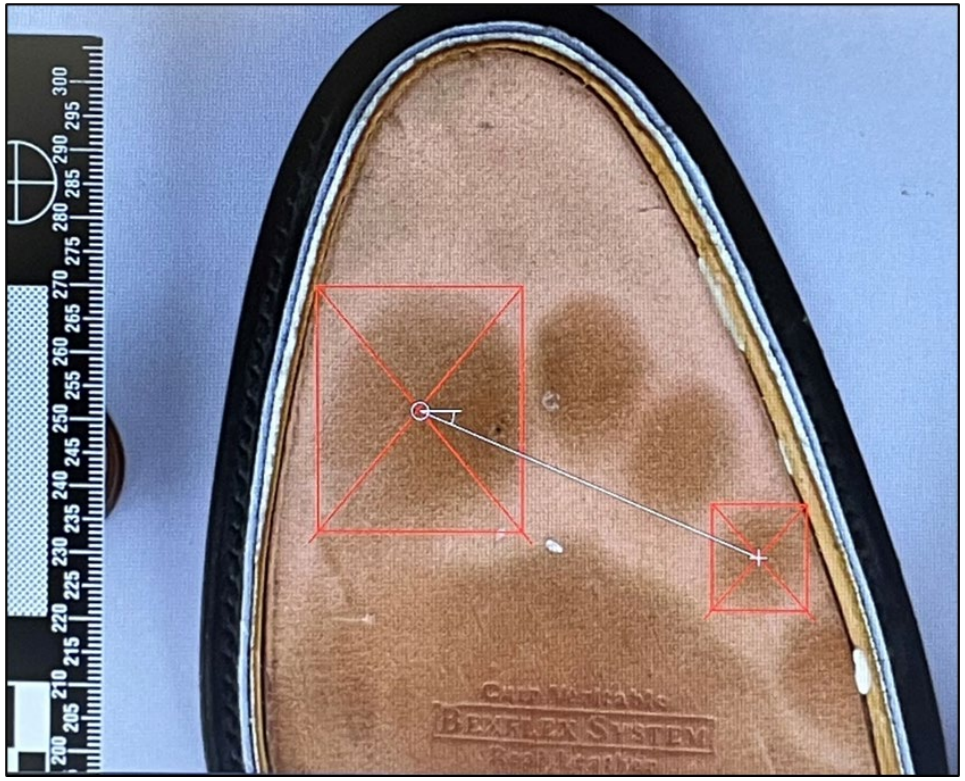
Method Development







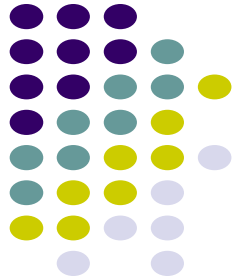
T1 to T4 measurement



Reliability Study

- Ethics Approval Granted
- 3 x Forensic Podiatrists
- **Sample:** 32 Insole Footprints
- Repeated Measures Design
- **Datasets:** First measurements & Repeat measurements
- **Units:** millimetres (mm)

Reliability Study



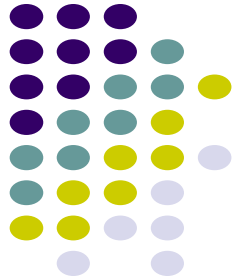
Sample Size:

- In principle when undertaking reliability studies, it is ideal to obtain at least 30 heterogeneous samples with at least 3 x raters (Koo & Li, 2016).
- From work of statistical power analysis by Cohen (1988) a sample of 30 would provide approximately 80% power.
- Minimum amount necessary for a repeated-measures design (Cohen, 1988, 2013).

Reliability Study

Descriptive Analysis

Measurement	Mean (mm)	Min (mm)	Max (mm)	Range (mm)	SD*	N
Rater 1						
T1 to T4 First	53.86	43.31	66.83	23.52	5.92	32
T1 to T4 Repeat	53.89	41.60	65.68	24.08	5.82	32
Rater 2						
T1 to T4 First	54.41	43.65	66.14	22.49	5.90	32
T1 to T4 Repeat	54.63	43.47	66.74	23.27	5.90	32
Rater 3						
T1 to T4 First	54.11	44.01	66.21	22.20	5.78	32
T1 to T4 Repeat	54.23	43.38	66.04	22.66	6.00	32



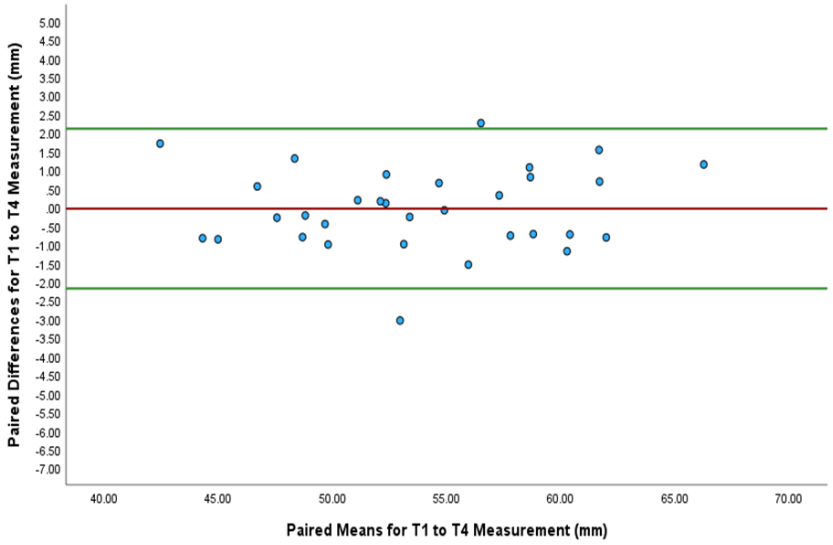
Repeatability
Intra Rater Analysis

Intra Rater Analysis

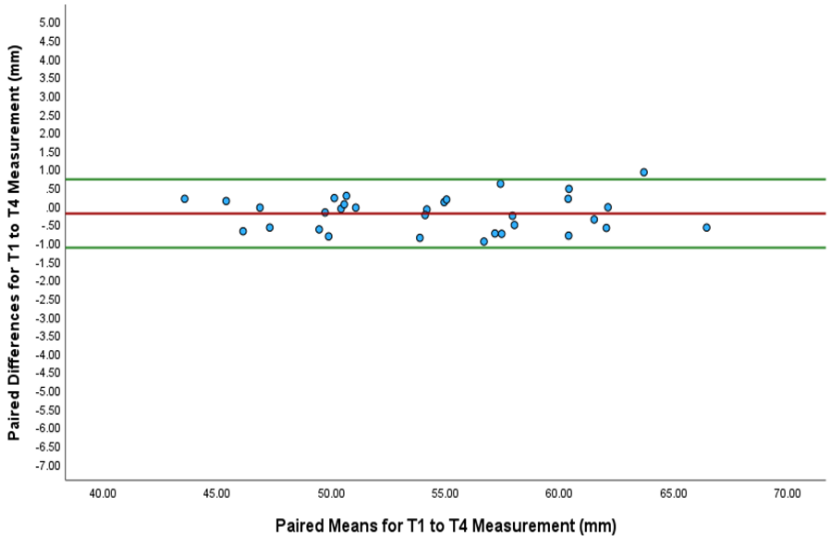
Measurement	SEM*	ICC**	95% CI		95% LOA***	
			Lower	Upper	Lower	Upper
Rater 1 First v Repeat	0.76	0.992	0.983	0.996	-2.18	2.10
Rater 2 First v Repeat	0.36	0.998	0.996	0.999	-1.14	0.71
Rater 3 First v Repeat	0.51	0.996	0.992	0.998	-1.53	1.29

Intra Rater Analysis LOA Plots

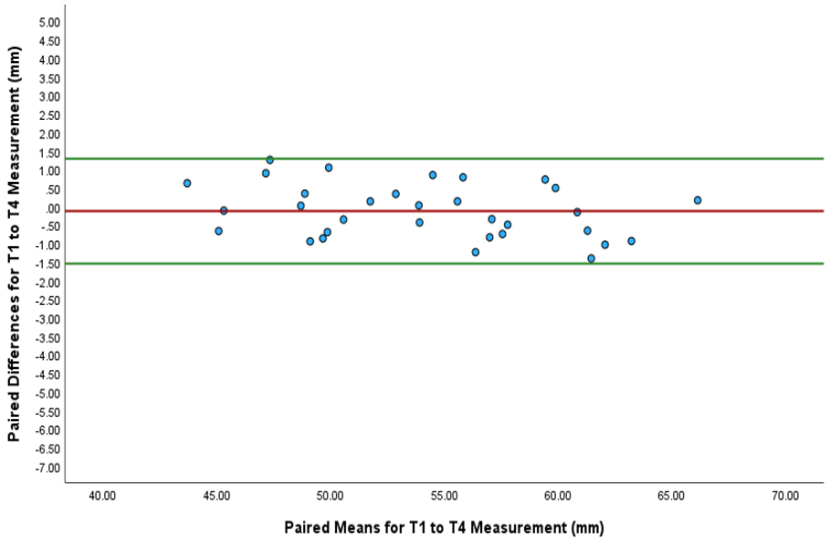
Rater 1

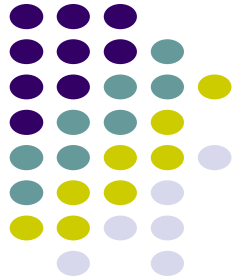


Rater 2



Rater 3





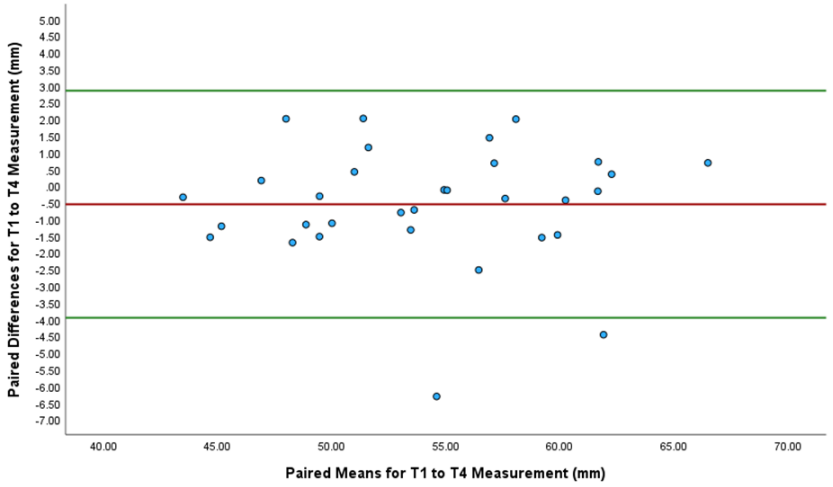
Reproducibility
Inter Rater Analysis

Inter Rater Analysis: FIRST

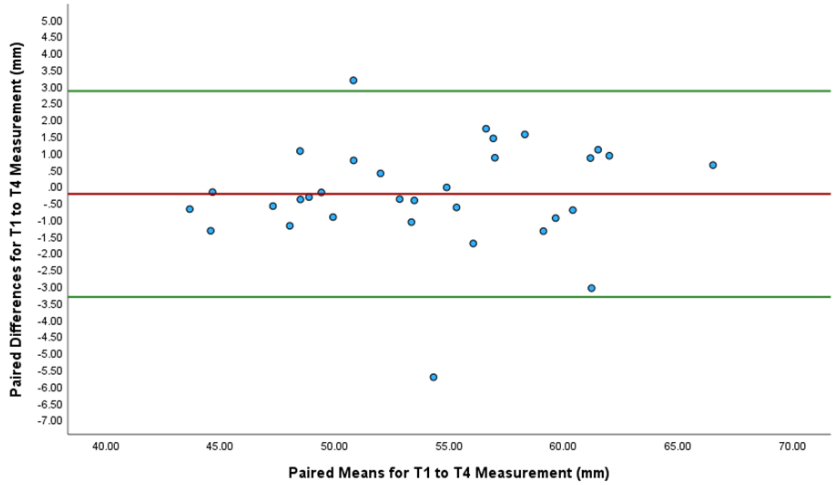
First Measurement	SEM*	ICC**	95% CI		95% LOA***	
			Lower	Upper	Lower	Upper
Rater 1 v Rater 2	1.26	0.976	0.952	0.988	-3.95	2.85
Rater 1 v Rater 3	1.10	0.982	0.963	0.991	-3.33	2.84
Rater 2 v Rater 3	0.48	0.997	0.993	0.998	-0.91	1.52

Inter Rater Analysis: LOA Plots FIRST

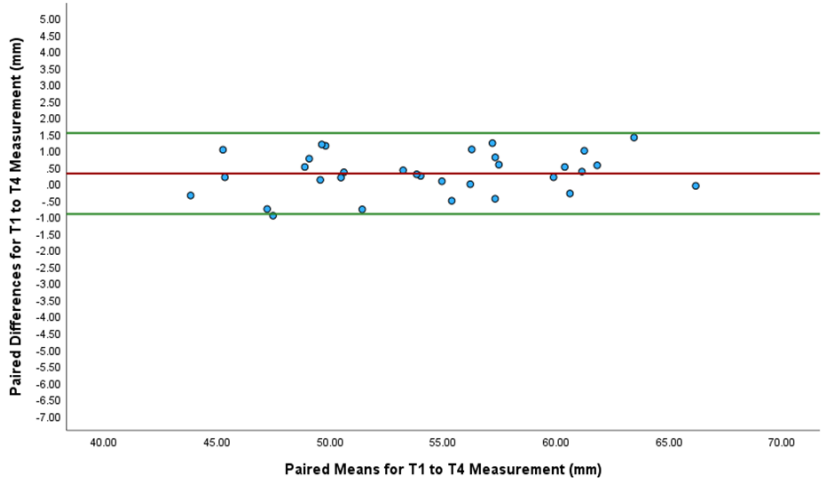
Rater 1 v Rater 2



Rater 1 v Rater 3



Rater 2 v Rater 3

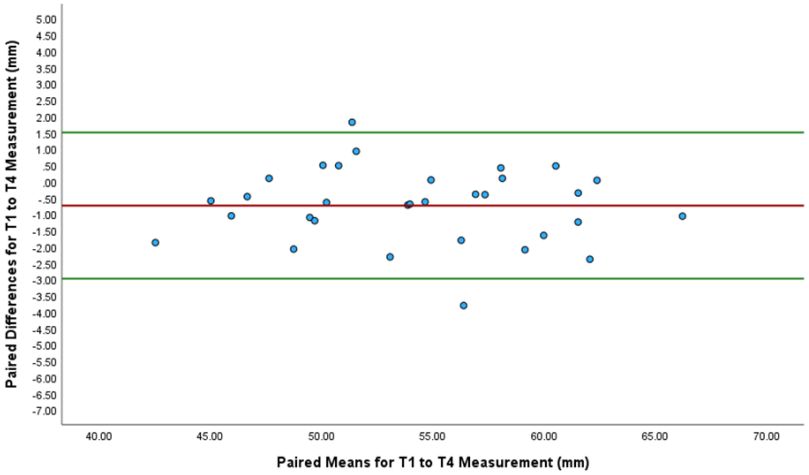


Inter Rater Analysis: REPEAT

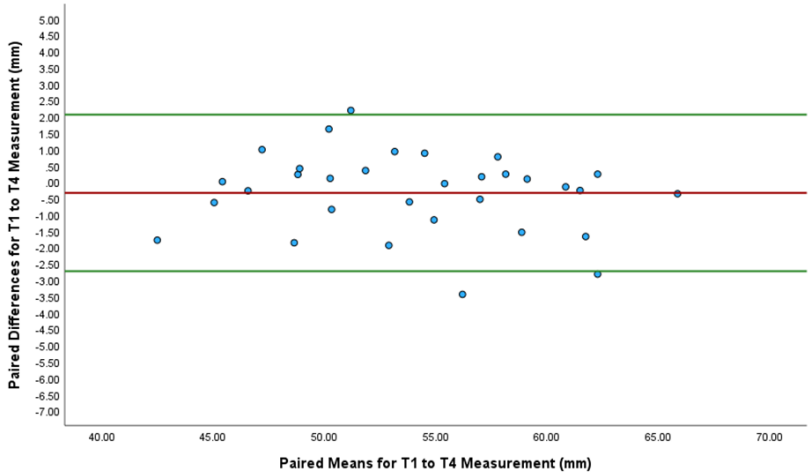
Repeat Measurement	SEM*	ICC**	95% CI		95% LOA***	
			Lower	Upper	Lower	Upper
Rater 1 v Rater 2	0.95	0.987	0.973	0.994	-2.97	1.50
Rater 1 v Rater 3	0.88	0.989	0.977	0.994	-2.72	2.06
Rater 2 v Rater 3	0.56	0.995	0.991	0.998	-0.98	1.78

Inter Rater Analysis: LOA Plots REPEAT

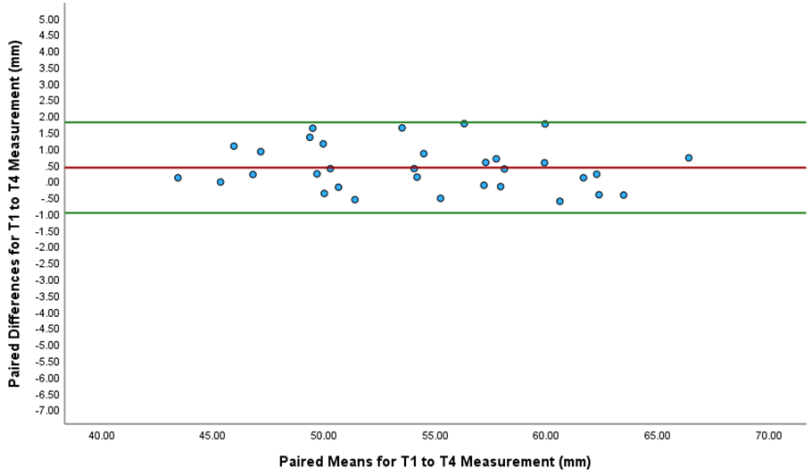
Rater 1 v Rater 2



Rater 1 v Rater 3

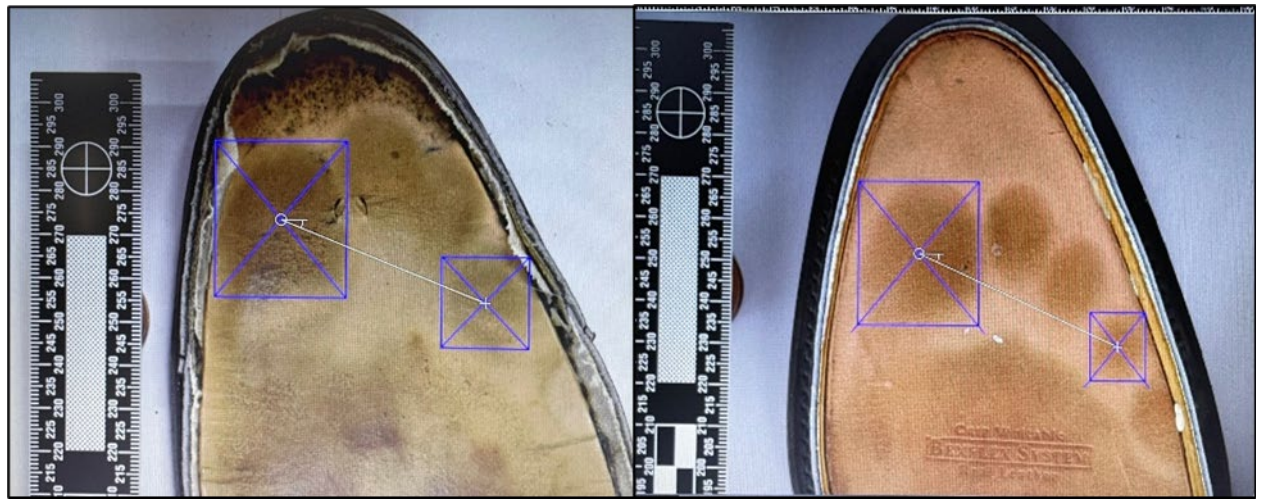


Rater 2 v Rater 3



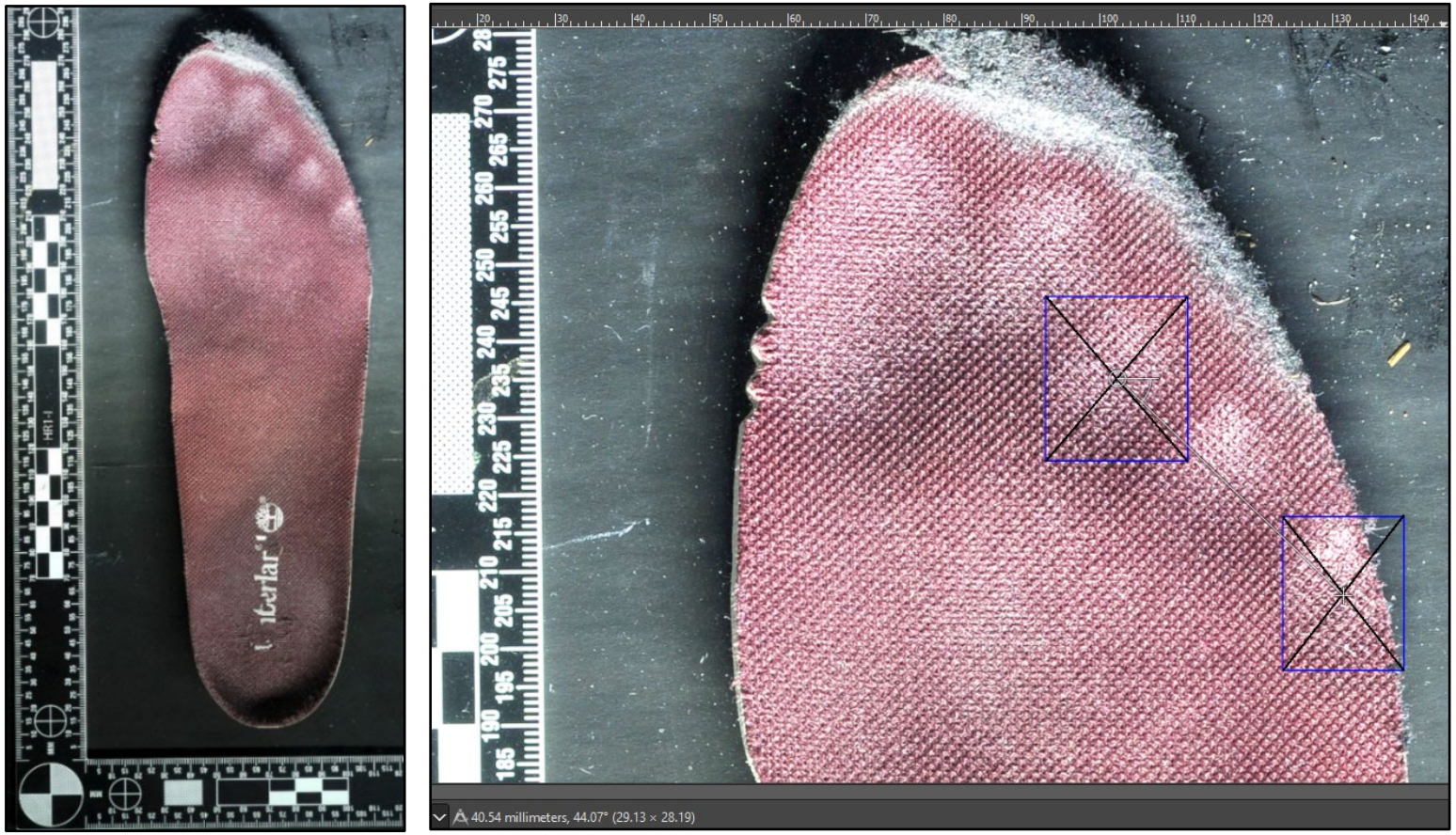
Conclusions

- High Intra-rater & Inter-rater Reliability
- High Levels of Consistency in Measurements Within & Between Raters
- Valid for Forensic Practice BUT with Limitations



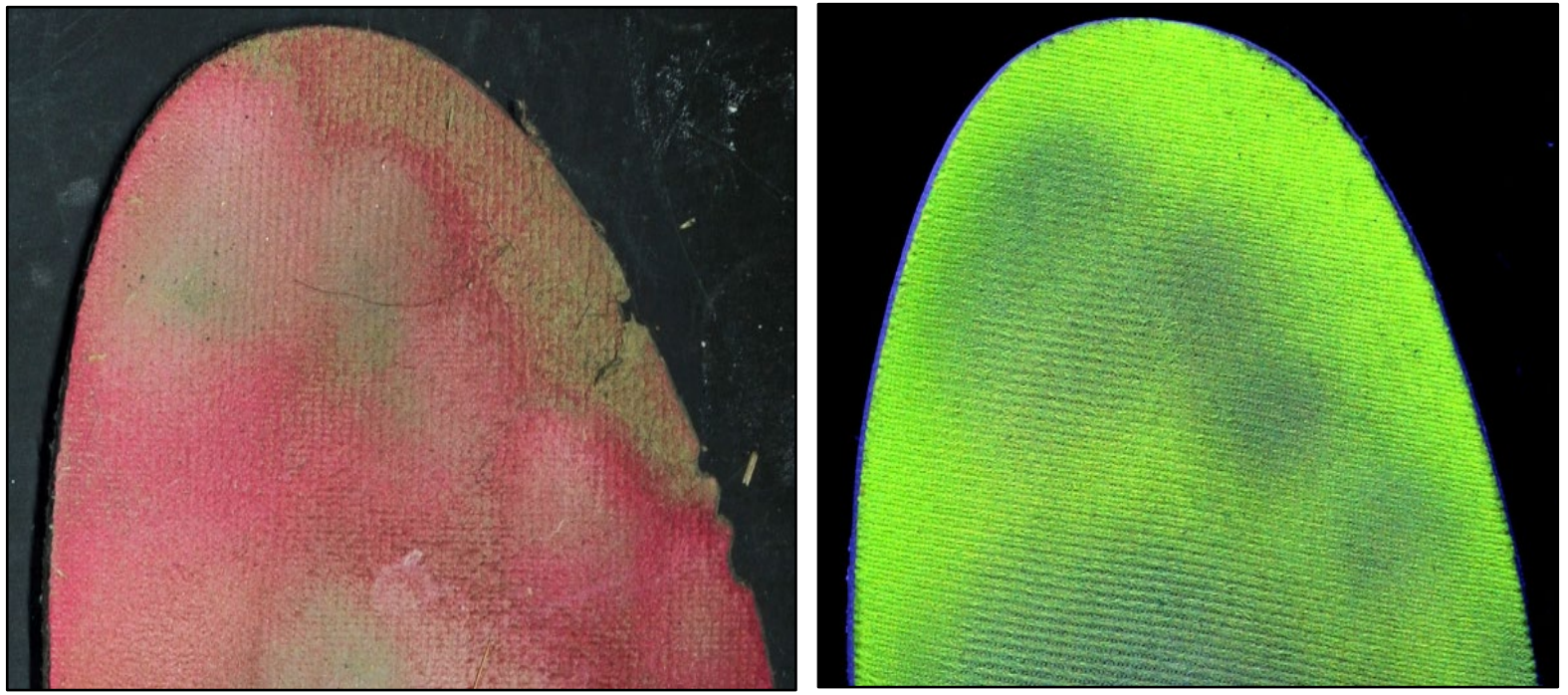
Limitations of HF Method

T5 - Toeprint



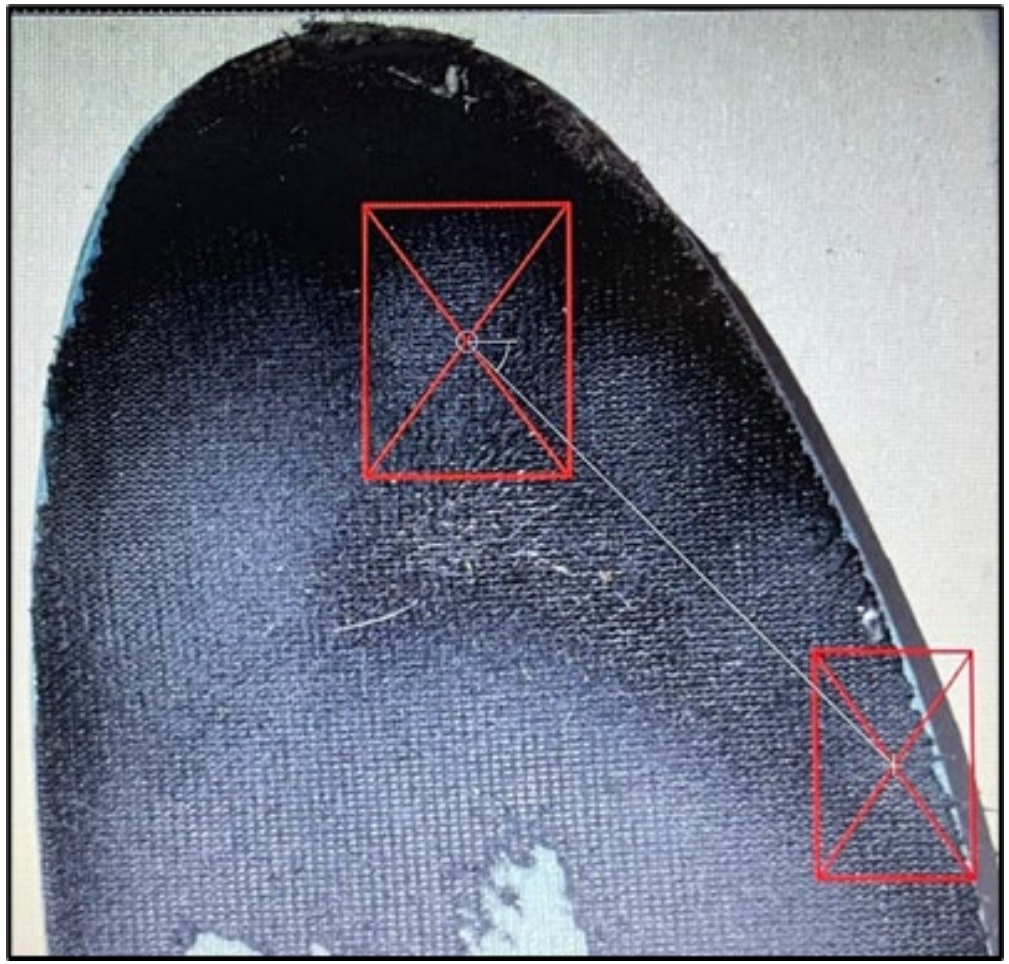
Limitations of HF Method

Coalescence / Merging of Toeprints



Limitations of HF Method

Toe Stem Print



References

Forensic Science Regulator (2020) Forensic Science Regulator Guidance Validation FSR-G-201, Issue 2, Crown Copyright, 1-60: https://assets.publishing.service.gov.uk/media/5f6b1a3de90e077ca292204f/201_-_FSR-G-201_Validation_Guidance_Issue_2.pdf

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15(2)

Portney, L. G., & Watkins, M. P. (2000). Foundations of clinical research: Applications to practice (1st ed.). Prentice Hall.

Reliability Studies:

Nirenberg, M. S., Ansert, E., & Reel, S. (2021). Reliability of a two-dimensional sock-clad footprint linear measurement method. Science & Justice, 61(5), 649-656

Reel, S., Rouse, S., Vernon, W., & Doherty, P. (2010). Reliability of a two-dimensional footprint measurement approach. *Science & Justice, 50(3),* 113-118

Q&A