Received: 12 March 2019 | Accepted: 28 April 2019

DOI: 10.1111/srt.12718

ORIGINAL ARTICLE

WILEY

The inter-rater and test-retest reliability of skin thickness and skin elasticity measurements by the DermaLab Combo in healthy participants

Kirsten Peperkamp[1] | Arico C. Verhulst[1,2] | Hanneke J. P. Tielemans[1] | Harm Winters[1] | Demi van Dalen[1] | Dietmar J. O. Ulrich[1]

[1]Department of Plastic, Reconstructive and Hand Surgery, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

[2]Radboudumc 3D Lab, Radboud University Medical Center, Nijmegen, The Netherlands

Correspondence
Kirsten Peperkamp, Geert Grooteplein Zuid 10, 6525 GA Nijmegen, The Netherlands.
Email: Kirsten.peperkamp@radboudumc.nl

Funding information
No funding was received in any form for the current work.

Abstract

Background: For several purposes, skin parameters like thickness and elasticity can be measured. However, little is known about the accuracy of those measurements.

Aim: The aim of this study was to determine the intrarater and test-retest reliability of skin thickness and elasticity measurements performed with the DermaLab Combo®.

Methods: A total of 49 participants were included in this study. Skin thickness and elasticity were measured at six defined locations on the dominant arm. Measurements were repeated two times by the same observer to determine the test-retest reliability. To determine the inter-rater reliability, a second observer repeated the measurements once.

Results: Inter-rater and test-retest reliability for elasticity measurements fluctuates per location and per parameter: Inter-rater intraclass correlation coefficient (ICC) ranged from 0.23 to 0.80, and test-retest ICC ranged from 0.25 to 0.84. Skin thickness was measured reliable by every observer on every location, with a test-retest ICC ranging from 0.71 to 0.83 and an inter-rater ICC ranging from 0.69 to 0.80.

Conclusion: The DermaLab Combo® showed a good inter-rater reliability when measuring skin thickness and elasticity. Not all locations are suitable for reliable inter-rater or test-retest measurements. The device is difficult to use by inexperienced users, as the echo probe is sensitive to small movements.

KEYWORDS

DermaLab Combo®, healthy participants, ICC, inter-rater, reliability, skin elasticity, skin thickness, test-retest

1 | INTRODUCTION

Measuring skin thickness and elasticity of patients can be of great value for several applications in clinical decision making.[1-4] A great amount of research have been performed over the years which relies on the accurate measurements of those parameters.[5,6] However, little is known about the reproducibility and accuracy of skin thickness and elasticity measurements on healthy subjects.

The DermaLab Combo® is able to measure multiple skin parameters, such as skin pH, elasticity, color, thickness, temperature and hydration. This article will focus on two of these parameters: skin thickness and skin elasticity. Skin thickness is determined by using

an echoprobe (20 MHz) which measures the thickness of the dermis and has a penetration capacity of 3.4 mm (Figure 1). The skin elasticity measurement is based on suction applied to the skin's surface, with a probe that adheres to the skin with a double adhesive sticker (Figure 2). This elasticity measurement takes both elevation and retraction into account. The device presents three parameters: ViscoElasticity (*VE*), Young's elasticity modulus (*E*), and skin retraction time (*R*). The DermaLab Combo® measures continuous variables.

Although there have been studies in literature related to the reliability of the DermaLab Combo® in people with burn scars, to the best of our knowledge no reliability studies have been performed on healthy individuals.[2-4]

The effect of a little increase or decrease of skin thickness has a big influence on the *VE* and *E* outcome, because the formula raises skin thickness to the third power. Most devices used in research, like the Cutometer from Courage & Khazaka or the skin fibrometer from Delfin Technology, do not take the effect of skin thickness on skin elasticity into account. Therefore, the skin elasticity outcome is not complete because skin thickness is an important parameter to display skin elasticity correctly.

Before using this device as a valid tool, it must be shown to be reliable and reproducible. Therefore, the aim of this study was to determine the intrarater and test-retest reliability of the DermaLab Combo® in a cohort of healthy participants.

2 | MATERIALS AND METHODS

This study was approved by the local ethics committee of the Radboud Medical Center. The protocols of the Declaration of Helsinki were taken into account. No funding or material support was provided for this study.

2.1 | Subjects

A group of 49 healthy participants—aged between 18 and 70—were included after giving written informed consent. Exclusion criteria were not giving written informed consent, skin disorders, lymphedema, pregnancy, tattoos on the upper extremities and the use of body lotion 24 hours before the measurements. The participants were asked not to use body lotion 24 hours before measurement because this affects the skin elasticity.[7]

2.2 | Instruments

Skin thickness and skin elasticity were assessed by the DermaLab Combo® (Cortex Technology).

Skin thickness—micrometer (μm)—thickness of the dermis determined by the intensity of the reflected ultrasound.

ViscoElasticity—*VE*—mega pascal (MPa)—calculation where both elevation phase and retraction phase are taken into account.

$$VE = \frac{E}{R\,\text{normalized}};\ \left(R\,\text{normalized} = \frac{R}{260}\right)$$

Young's elasticity modulus—*E*—mega pascal (MPa)—the force necessary to elevate the skin surface 1.5 mm, followed by a calculation in which the skin thickness is taken into account.

$$\Delta x = \varphi \cdot p \frac{r^4}{E \cdot S^3}$$

Δx = deviation, middle of surface; φ = constant; p = surface pressure; E = elasticity modulus; R = radius of the surface; S = skin thickness.

Skin retraction time—*R*—seconds (S)—time the skin takes to retract 1.5 mm after elevation.

2.3 | Study protocol and procedure

The skin elasticity was measured twice at six defined points on the dominant arm by the first observer (R1). A second observer (R2) performed the measurements at the same defined points. Both observers had received training in using the DermaLab Combo®. Measurements were performed with an interval of

FIGURE 1 20-MHz echoprobe to measure skin thickness, DermaLab Combo® [Colour figure can be viewed at wileyonlinelibrary.com]

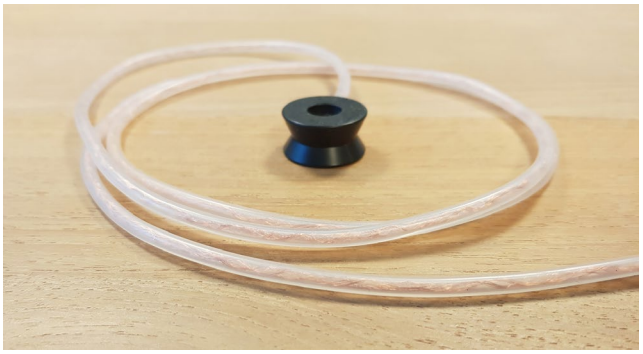

FIGURE 2 Suction device to measure skin elasticity, DermaLab Combo® [Colour figure can be viewed at wileyonlinelibrary.com]

45 minutes between each measurement cycle. Both observers measured independently and were blinded to the results of each other. Measurements took place from January 2018 to March 2018. The six defined points were chosen because of the relative easy accessibility. Figure 3 shows the location of the points 1, 2, 3, 4, 5, and 6. The temperature and humidity were recorded during every measurement. All measurements were performed with the participant in a sitting position with the dominant arm in a 90-degree angle relaxed on a table. While measuring, body movement was avoided. If the measurement is suspected wrong, it has been repeated.

First, six points on the arm were signed with a black marker, so they would remain visible during all measurements. Both observers assessed these six marked areas using the DermaLab Combo®. Elasticity measurements performed with the DermaLab Combo consist of two parts: First, the skin thickness is measured by using ultrasound. Next, the elasticity is measured with a suction device.

After an acclimatization period of 30 minutes, recommended by the manufacturer, the ultrasonography was performed by R1 at each marked point. Next, R1 measured the elasticity of the skin at each marked point. After this set of examinations, the skin must be given a rest period of at least 45 minutes. During this waiting period, a questionnaire was completed with additional information about gender, age, weight, length, and use of body lotion.

Subsequently, the second set of measurements were performed by R2. Likewise, the skin thickness at each marked point was determined and written down followed by another elasticity measurement on the same points.

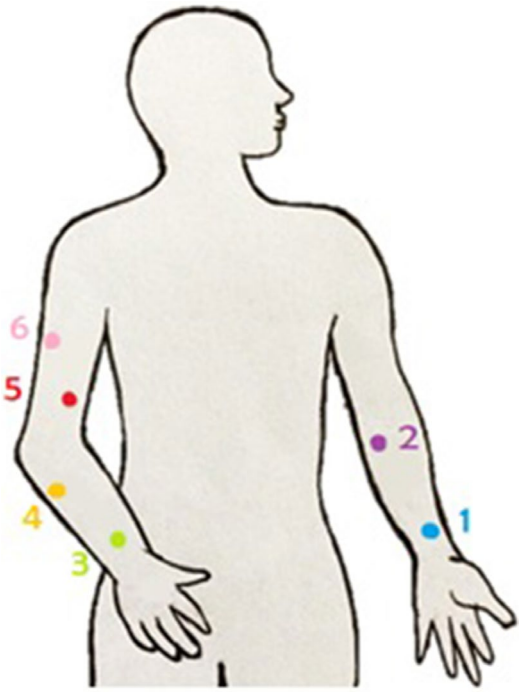

FIGURE 3 Location of marked points, numbered from 1-6. For each participant, the dominant arm was measured [Colour figure can be viewed at wileyonlinelibrary.com]

Before the third measurement was started by R1, another 45-minute pause was scheduled to give the skin some rest. Afterward, the last six skin thickness and elasticity measurements took place.

This whole process is schematically displayed in Figure 4.

2.4 | Statistical analysis

Data analysis was performed by a biostatistician, using SPSS version 22 (IBM) and SAS/STAT (SAS Institute Inc).

The participants' characteristics were described using descriptive statistics.

Reliability refers to the extent to which measurements can be replicated, that is, whether measurement is the same after repeated trials.[8] Inter-rater reliability is the extent of agreement among raters scoring the same participants under the same conditions.[9] Test-retest reliability refers to the reliability or the consistency of a measurement over time.[10]

Inter-rater reliability was determined from two measurements done by two observers. Test-retest reliability was determined from two measurements done by the same observer. Both reliability parameters were analyzed in the same way. Using the ICC (two-way mixed effects), SD, and a mixed model analysis (PROC MIXED by SAS), reliability was assessed. Random effects and covariances are implemented in this statistical model. This analysis can compute efficient estimates of fixed effects.[11] This means there has been corrected for effects of the observer and the measured location, since these factors influence the outcome.

All ICC were interpreted using the Rosner interpretation (0-0.40: poor agreement; >0.40-0.75: good agreement; >0.75-1.00: excellent agreement).[12] The higher the ICC, the stronger the correlation. A high correlation means that the measurement can be performed reliable by different observers or can be performed reliable by the same person multiple times.

3 | RESULTS

Table 1 shows the ICC's and SD's of this research. A total of 49 participants were included in this study (23 [46.9%] men, 26 [53.1%] woman, mean age 38.7 [range 20-69]). The mean BMI was 24.9 (SD 3.07). Of all participants, 42 (85.7%) were right-handed.

The values for skin thickness in healthy skin ranged from 705 to 2120 μm (mean 1308, SD 254).

The values for *VE* in healthy skin ranged from 0.2 to 15.1 MPa (mean 2.9, SD 2.0).

The values for *E* in healthy skin ranged from 0.3 to 11.5 MPa (mean 2.4, SD 1.43).

The values for *R* in healthy skin ranged from 260 to 2989 s (mean 49.4, SD 235.5).

The intraclass correlation coefficient (ICC) and standard deviation (SD) were determined for each measured point on the dominant arm.

TABLE 1 ICCs and SDs for the intrarater and test-retest reliability of the skin thickness, *VE*, *E*, and *R* parameter measured with the DermaLab Combo®, taken on healthy skin

	Mean	Interobserver SD	Interobserver ICC	Test-retest SD	Test-retest ICC
Skin thickness					
Location					
1	1330.13	104.0	0.80[a]	207.1	0.83[a]
2	1089.09	84.2	0.76[a]	151.4	0.77[a]
3	1435.85	106.6	0.73[a]	174.5	0.81[a]
4	1270.84	110.4	0.74[a]	186.2	0.74[a]
5	1257.59	116.8	0.71[a]	180.8	0.71[a]
6	1469.34	150.5	0.69[a]	226.8	0.74[a]
VE					
Location					
1	2.208	0.9	0.47[a]	0.8	0.66[a]
2	4.996	1.5	0.60[a]	1.8	0.62[a]
3	1.622	0.6	0.57[a]	0.7	0.57[a]
4	3.041	1.1	0.70[a]	1.6	0.70[a]
5	3.166	1.1	0.64[a]	1.5	0.65[a]
6	2.224	1.0	0.45[a]	0.9	0.56[a]
E					
Location					
1	2.141	0.5	0.74[a]	0.9	0.84[a]
2	3.877	1.1	0.60[a]	1.3	0.63[a]
3	1.626	0.4	0.64[a]	0.5	0.75[a]
4	2.657	0.7	0.70[a]	1.1	0.73[a]
5	2.579	0.6	0.69[a]	1.0	0.69[a]
6	1.635	0.7	0.46[a]	0.7	0.66[a]
R					
Location					
1	554.10	259.2	0.23[a]	143.2	0.25[a]
2	407.62	61.0	0.68[a]	89.8	0.68[a]
3	611.50	198.2	0.58[a]	234.3	0.59[a]
4	534.03	174.8	0.50[a]	175.7	0.56[a]
5	473.06	120.0	0.54[a]	129.7	0.58[a]
6	392.20	45.3	0.76[a]	79.7	0.79[a]

[a]Significant at 0.01 level (2-tailed).

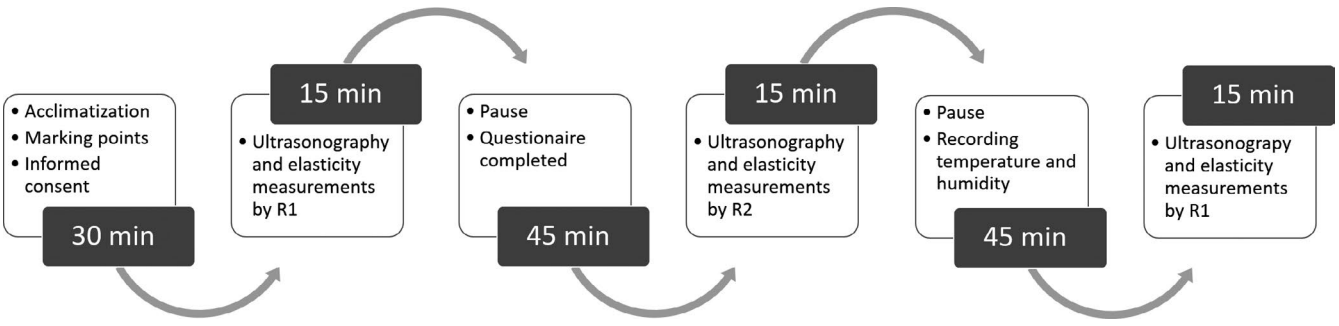

FIGURE 4 Schematically displayed process

3.1 | Inter-rater reliability

The ICC for determining inter-rater reliability for **skin thickness** showed a good to excellent agreement which ranged from 0.69 (location 6) to 0.80 (location 1).

The ICC for determining the inter-rater reliability for *VE* showed a good agreement which ranged from 0.45 to 0.70. Location 4 has the highest ICC for inter-rater reliability, whereas location 1 and 6 showed the lowest ICC.

The ICC for determining the inter-rater reliability for ***E*** showed a good agreement which ranged from 0.46 (location 6) to 0.74 (location 1).

The ICC for determining the inter-rater reliability for *R* showed a poor to excellent agreement which ranged from 0.23 (location 1) to 0.76 (location 6). All locations showed a good agreement using the Rosner interpretation, except location 1 which ICC deviates from the other measured locations.

3.2 | Test-retest reliability

The ICC for determining the test-retest reliability for **skin thickness** showed a good to excellent agreement which ranged from 0.71 to 0.83. Location 1 and 3 had the highest ICC for test-retest reliability, whereas location 5 showed the lowest ICC.

The ICC for determining the test-retest reliability for *VE* showed a good agreement which ranged from 0.56 to 0.70. Location 4 has the highest ICC for test-retest reliability, whereas location 3 and 6 showed the lowest ICC.

The ICC for determining the test-retest reliability for *E* showed a good to excellent agreement which ranged from 0.63 (location 2) to 0.84 (location 1).

The ICC for determining the test-retest reliability for *R* showed a poor to excellent agreement which ranged from 0.25 (location 1) to 0.79 (location 6). All locations showed a good or even excellent agreement using the Rosner interpretation, except location 1 which ICC deviates from the other measured locations.

4 | DISCUSSION

This study investigated the reproducibility of skin thickness and elasticity measurements performed with the DermaLab Combo®. The results of this study demonstrated that this device has a wide range of reliability outcomes when performing skin thickness and elasticity measurements on healthy subjects. Most measured locations demonstrated a good inter-rater reliability for all parameters. This means different people can repeat measurements reliable. Yet it is notable that both location 1 and 6 showed extremely variable inter-rater reliability. Test-retest reliability varied greatly over the different measured locations. Therefore, based on this study, it is not recommended to repeat measurements by the same person.

The test-retest and inter-rater reliability for determining skin thickness was high for all measured locations. The reliability measurements for elasticity were high in particular on the forearm. The reliability outcomes for elasticity were variable on the upper arm, indicating that not all locations are suitable for reliable inter-rater or test-retest measurements. All measured locations are suitable for a repeated, reliable skin thickness measurement. The elasticity parameters (*VE*, *E*, and *R*) have shown a wide variety of outcomes at the various measured locations. As a result, it is not possible to identify an unambiguous location that can be measured reliable repeatedly. This is possibly caused by differences in skin thickness.

This is, to our knowledge, the first study to assess the accuracy of the DermaLab Combo. Previously, this device has only been validated in burn scars research. Gankande et al[2] performed scar assessment using two methods (DermaLab Combo® and modified Vancouver Scar Scale [mVSS] score) in 100 subjects to determine the validity of the DermaLab Combo. Gankande et al[3] continued researching the application of the DermaLab Combo® in burn scars by determining the inter-rater and test-retest reliability in 30 patients with burn scars. Inter-rater and test-retest reliability was found to be excellent for pigmentation, pliability, and thickness. For determination of vascularity, the inter-rater reliability was excellent, but test-retest reliability was low. Anthonissen et al[4] investigated the reproducibility of repeated elasticity and transepidermal water loss measurements with the DermaLab Combo® in 32 burn scars patients, where results show a good reliability. They recommend to use a mean of repeated measurements to be able to interpret the results correctly.

Hua et al[13] investigated the differences and correlation between the outcomes of two devices (DermaLab Combo® vs detectors from Courage and Khazaka) by measuring 30 healthy subjects. The outcomes regarding skin elasticity were statistically negative correlated. Inter-rater reliability was not determined. Test-retest reliability (in the article mentioned as repeatability) was high.

In our study, skin thickness and elasticity measurements of the skin were performed at four locations on the inner and outer forearm and two locations on the upper arm, on the dorsal and ventral side, respectively (Figure 3). These six locations were chosen for a balanced distribution of measurement points. The upper limb is easy accessible, without any need to undress. To measure consistently, we measured the dominant arm in each subject. Because no power analysis could be performed, we opted for a research group size of 50 people. One person was excluded from this study because of using body lotion 24 hours before measurement.

During the measurements, certain practical limitations became visible. The plastic water barrier film on the echo probe had to be replaced often because of damage. While the instruction manual demonstrated how to do this, it proved difficult to properly replace the film without trapping air underneath, which could cause measurement errors. Therefore, a skilled and trained researcher has to be available to perform these measurements.

Most of the time, the skin thickness could not be determined after creating a single ultrasound image; therefore, several attempts were needed before a correct measurement could be obtained. Furthermore, the echo probe was extremely sensitive

to small movements, pressure, and the amount of ultrasound gel during the measurements. Small differences between the observers handling the probe led to different values. For example, small movements could cause a large difference in measured skin thickness. Because V and *VE* are both reliant on the skin thickness, this might as well explain the difference between the test-retest and inter-rater reliability.

In this study, the dominant arm, the most trained arm, was measured in every participant. The results and conclusion are based on six measurements of the arm. The outcomes might be different if the device is applied to other parts of the body.

The study protocol provided a minimum time period of 45 minutes between consecutive measurements in order to avoid consequences of hysteresis. The measurements took place at the participant's own home, because of logistic reasons. The advantage of measuring the participant in their own environment is that the skin is already acclimatized to the surrounding. The conditions in which the measurements have taken place correspond to the hospital environment in which measurements normally will take place.

The results of this study indicate that this device can be used for reliable inter-rater measurements of skin thickness and elasticity of the arm. This information can be used to interpret the results of future research regarding elasticity of the skin. In addition, the DermaLab Combo® provides continuous data that may be useful to monitor skin changes over time. Follow-up research will be necessary to investigate the normal changes of the skin over a longer period to optimize the interpretation of the outcomes. The DermaLab Combo® might also be useful for multiple other purposes like determining the effect of lymphedema treatment, burning scar assessment, and monitoring wound healing.

5 | CONCLUSION

This study showed a wide range of all outcomes measured in healthy subjects. Most locations showed a good to excellent reliability for inter-rater reliability. In conclusion, this means two measurements can be repeated reliable by two different researchers. However, based on this study, it is not recommended to repeat measurements by the same person. The echo probe is sensitive to movements, so it is recommended a trained user carries the skin thickness measurement out.

CONFLICT OF INTEREST

The authors state no conflict of interest.

ORCID

Kirsten Peperkamp https://orcid.org/0000-0002-2511-3840

REFERENCES

1. Killaars RC, Penha TR, Heuts EM, van der Hulst RR, Piatkowski AA. Biomechanical properties of the skin in patients with breast cancer-related lymphedema compared to healthy individuals. *Lymphat Res Biol*. 2015;13(3):215-221.
2. Gankande TU, Duke JM, Wood FM, Wallace HJ. Interpretation of the DermaLab Combo® pigmentation and vascularity measurements in burn scars assessment: an exploratory analysis. *Burns*. 2015;41(6):1176-1185.
3. Gankande TU, Duke JM, Danielsen PL, DeJong HM, Wood FM, Wallace HJ. Reliability of scar assessments performed with an integrated skin testing device – the DermaLab Combo ®. *Burns*. 2014;40(8):1521-1529.
4. Anthonissen M, Daly D, Fieuws S, et al. Measurement of elasticity and transepidermal water loss rate of burn scars with the Dermalab®. *Burns*. 2013;39(3):420-428.
5. Schlangen L, Brokken D, van Kemenade PM. Correlation between small aperture skin suction parameters: statistical analysis and mechanical model. *Skin Res Technol*. 2003;9(2):122-130.
6. Sulaiman W, Azad AK, Daud NA, Khan N. Evaluation of skin elasticity after used different seaweed containing products by DermaLab® Combo. *Asian J Pharm Pharmacol*. 2016;3(2):72-76.
7. Hadi H, Awadh AI, Hanif M. The investigation of the skin biophysical measurements focusing on daily activities, skin care habits, and gender. *Skin Res Technol*. 2015;2016(22):247-254.
8. Eriksson P, Kovalainen A. *Qualitative Methods in Business Research*, 2nd edn. London, UK: Sage; 2015.
9. Li Gwet K. *Handbook of Inter-Rater Reliability*, 4th edn. Gaithersburg, MD: Advanced Analytics LLC; 2014.
10. Kline PA. *Handbook of Test Construction: Introduction to Psychometric Design*. New York, NY: Methuen; 1986.
11. Littell RC, Henry PR, Ammerman CB. Statistical analysis of repeated measures data using SAS procedures. *J Anim Sci*. 1998;76(4):1216-1231.
12. Rosner BA. *Fundamentals of Biostatistics*, 7th edn. Boston, MA: Brooks/Cole Cengage learning; 2010.
13. Hua W, Fan L-M, Dai R, et al. Comparison of two series of non-invasive instruments used for skin physiological properties measurements: the DermaLab® from Cortex Technology vs. the series of detectors from Courage & Khazaka. *Skin Res Technol*. 2017;23(1):70-78.

How to cite this article: Peperkamp K, Verhulst AC, Tielemans HJP, Winters H, van Dalen D, Ulrich DJO. The inter-rater and test-retest reliability of skin thickness and skin elasticity measurements by the DermaLab Combo in healthy participants. *Skin Res Technol*. 2019;25:787–792. https://doi.org/10.1111/srt.12718