# Intra-rater reliability and agreement of the Indurometer when used to assess mid-calf tissue compressibility among people affected by moderate to severe lymphedema in Bangladesh and Ethiopia.

Running title: Indurometer reliability in Ethiopia and Bangladesh

Janet Douglass PhD1\*, Hayley Mableson PhD1, Sarah Martindale MPH1, Jahirul Karim2, ASM Sultan Mahmood MPH2, Fikre Hailekiros MPH3, Louise Kelly-Hope PhD1

1. Liverpool School of Tropical Medicine, Department of Tropical Disease Biology, Centre for Neglected Tropical Diseases.
2. Bangladesh Ministry of Health and Family Welfare, Filariasis Elimination Program.
3. National Podoconiosis Action Network.

\* Corresponding author, [jan.douglass@lstmed.ac.uk](mailto:jan.douglass@lstmed.ac.uk)

## Abstract

**Background.** Lymphedema is a chronic skin disease which induces tissue fibrosis (stiffness). Tissue tonometry has been used to assess skin and tissue compressibility in lymphedema, primarily in research on arm lymphedema after breast cancer. A digital tonometer, the Indurometer (Flinders and SA Biomedical Engineering, Australia), has shown excellent intra-rater reliability in young healthy people in Australia and Myanmar and was able to detect covert changes in young, asymptomatic Myanmar people infected with lymphatic filariasis. It has not been tested in overt lower-limb lymphedema. **Methods and results.** The Indurometer was used in a cluster randomized trial in Bangladesh and Ethiopia to measure tissue compressibility among adults affected by moderate to severe lymphatic filariasis- or podoconiosis-related leg lymphedema. The study compared different self-care intervention and after baseline there were follow-ups at 4-, 12- and 24-weeks. Three consecutive Indurometer scores were collected by two data collection teams in each country at the mid-calf of each leg. Indurometer scores were available for three time-points in Bangladesh and four time-points in Ethiopia. An intraclass correlation coefficient (ICC) was calculated for each data collection team and a coefficient of variation (CV) was used to assess measurement agreement in moderate and severe stages of lymphedema. The intra-rater reliability among local research assistants was good to excellent in both countries at all time-points (ICC range 0.829 (95% CI 0.730, 0.896) - 0.992 (95% CI 0.989, 0.995)). In Bangladesh agreement between measures was highest among unaffected legs (range 16% - 22%) and lowest in severe lymphedema (range 19% - 39%). CV scores in Ethiopia showed no distinct pattern for lymphedema stage (range 15% - 32%). **Conclusion**. The Indurometer is an inexpensive and easy to use device to assess skin and tissue compressibility and should be considered in clinical research on lower-limb lymphedema.

Condensed abstract

Changes to the skin and subcutaneous tissue in lymphedema alter the compressibility (stiffness or suppleness) of the affected tissues. This change in tissue compressibility can be measured using tissue tonometry and several devices have been used previously in research on arm lymphedema. A hand-held electro-mechanical tonometer, the Indurometer, was used in a study in Bangladesh and Ethiopia on people with lymphatic filariasis- or podoconiosis-related leg lymphedema. Indurometer scores from two research teams in each country were analysed and showed good to excellent intra-rater reliability, which means that each time three consecutive measures were taken the variation between them was very small. The Indurometer is a field-friendly device that requires minimal training to operate and is suitable for use in resource poor settings.

## Introduction

Lymphedema is a chronic skin disease affecting between 100 and 250 million people world-wide and there are multiple causes including some oncology treatments, exposure to irritant soils, congenital malformation and parasitic infection [1-3]. Skin and subcutaneous tissue pathology typically progress through identifiable phases, or stages, regardless of the initiating aetiology [4], and while there is no universally accepted criteria for clinical staging [5], at its simplest lymphedema can be categorized into mild, moderate and severe. Mild stages are characterised by a protein rich subdermal swelling which may fluctuate and reduce with elevation or overnight. In the moderate or middle stages, fatty deposits and fibrous tissue invade the subcutaneous compartment which is now permanently enlarged despite elevation, and the skin begins to form into permanent folds. Severe or advanced stages are characterised by visible changes in the skin including hyperkeratosis, deep folds and papillomatosis (warty growths). Comprehensive assessment of lymphedema status may involve sophisticated imaging such as lymphoscintigraphy or very simple tape measures of limb circumference, and a combination of measures is typically employed to capture various aspects of the pathology [5]. In resource poor settings limb size is frequently used as the primary outcome measure but size alone does not capture changes in the thickness of the skin or the fibrous state of the underlying tissue, both of which can improve or deteriorate without a change in limb size [6].

Tissue tonometry was developed as a measure of fibrosis in lymphedema in 1976 [7]. Although first used in 1992 to measure fibroplasia in lower-limb LF-related lymphedema [8], it has primarily been used in research on breast cancer-related lymphedema (BCRL) of the arm [9-13], and there are few studies on tissue tonometry in leg lymphedema [14, 15]. Tissue tonometers employ a small indenter pressed into the skin under a defined load to determine the compressibility or stiffness of the tissue, and first-generation mechanical tonometers have been superseded by digital, or electro-mechanical versions [16, 17]. One such electro-mechanical device, the Indurometer (Flinders and SA Biomedical Engineering, Australia), has previously demonstrated excellent intra-rater reliability when used on the legs of healthy young people in Australia and Myanmar [18] It was further used in a study on asymptomatic young people at risk of LF-related lymphedema in Myanmar [19], but no reliability analysis has been performed on its use among people affected by overt lower-limb lymphedema.

The Indurometer was used to measure tissue compressibility at the calf in a study on adults affected by lower-limb lymphedema in Bangladesh and Ethiopia. All data was collected by local research assistants recruited within in each country, and an intra-class correlation coefficient for each data collector is reported here. The Indurometer has been shown to be reliable in multiple populations and whilst factors which are known to moderate Indurometer scores such as hydration [20] should always be considered, operator variability can be expected to be low.

## Materials and Methods

The study was a cluster randomized trial on people affected by lymphedema in rural areas of Bangladesh and Ethiopia, as part of a longitudinal study on community-based home-care. Indurometer scores were collected on people aged over 18 years of age with moderate to severe lymphedema and with no co-morbidities who had been recruited through local health centers. In Bangladesh, personnel from the Ministry of Health Family and Welfare (MOHFW) and the Centre for Injury Prevention and Research, Bangladesh (CIPRB) in Dhaka formed the research teams. In Ethiopia, personnel from the Federal Ministry of Health (FMOH) and local health staff in Simada *Woreda* (district) formed the research teams. There were two teams in each country and all teams received three days of training in the data collection activities, including practice in use of the Indurometer. Each of the four data collection teams had a designated clinical officer who operated the Indurometer and collected the measures, and a designated data collector who entered the scores to the Open Data Kit Collect (ODK Collect)[21] application loaded to an electronic tablet (Samsung Galaxy Tab A 10.1). Indurometer measures were collected at baseline and at three follow-ups at 4-weeks, 12-weeks and 24-weeks.

The Guidelines for Reporting Reliability and Agreement Studies (GRRAS) guided analysis and interpretation of the results on intra-rater reliability and agreement of the Indurometer [22] . The study was approved by the Liverpool School of Tropical Medicine Research Ethics Committee (approval number 012-18), the Bangladesh Human Medical Research Committee, and the Amhara Public Health Institute Research Ethics Committee. The study was registered on the ISRCTN Registry, trial number 16764792. All subjects gave their informed consent for inclusion before they participated in the study and the study was conducted in accordance with the Declaration of Helsinki [23].

### Indurometer measures

The ‘Indurometer’ (BME-1563, Flinders and SA Biomedical Engineering, Australia) is designed to quantify the compressibility of flexible material such as the skin and subcutaneous tissue. It is a battery-operated, hand-held unit a with a 7-cm diameter Perspex® reference plate through which a 1-cm diameter plunger (indenter) can extend. Images of the Indurometer can be seen in Figure 1. To take the measure the device is aligned with the reference plate close to, and parallel with the skin over the measurement point, and the operator smoothly and evenly presses the device into the skin. As the pressure increases, the reference plate remains level with the surrounding skin while the indenter can continue pressing into the skin and underlying tissue. A force sensor detects when a load equivalent to 200-g has been applied to the indenter and the device emits a single beep to indicate the measure has been completed successfully. The device is removed from the skin and the result can be read from a light emitting diode (LED) screen.

The result is delivered numerically with two decimal places and represents the distance in millimetres to which the indenter has extended (was able to compress the skin and tissue) beyond the reference plate. A higher score is given when the indenter has extended further into the tissue, indicating more tissue compressibility (less tissue stiffness). Tissue that is less compressible (more tissue stiffness) will return lower scores as the indenter will not be able to press in as far under the same 200-g load. A score of 0.00 represents no movement of the indenter beyond the level of the reference plate. Scores approaching double digits indicate very soft, easily compressible tissue. Incorrectly performed measures, such as those that are too fast, too slow or erratic in movement are identified with a double beep and an error message appears on the LED screen. The device must then be reset, and the measure repeated. The procedure is non-invasive and should cause no discomfort to the patient. The Indurometer has a self-calibrating procedure which was performed daily before the first measure. Images of the instruction sheets provided to the data collectors for device operation can be viewed in Supplementary Figure 1a) operation and b) recalibration.

### Data collection

Measures were taken on each leg at mid-calf. To find the mid-point the patient was asked to lie prone on a treatment bench if one was available, or a yoga mat on the floor if not. A second yoga mat was used to support the ankles and ensure the leg could be held in a relaxed state with the ankle in a neutral position. The clinical officer used a retractable tape measure to determine the length of the calf from the most obvious crease behind the knee to the base of the heel. The tape was held with the zero mark at the crease behind the knee and if a single major crease was not obvious the lower leg was raised and the crease that formed was marked with a washable skin marker. The length measure was halved to find the mid-point and with the zero mark of the tape measure still held at the crease behind the knee, a small line was made with the skin marker on either side of tape at the mid-point, so that when the tape was removed there was a horizontal line with a gap in the middle across the middle of the calf. The indenter of the Indurometer was aligned with the gap to perform the measures. Both legs were measured beginning with the left leg and all three measure were taken sequentially with only enough time between each test to reset and reposition the device. In both countries the Indurometer measure was collected by the clinical officer in each team and the data collector visually confirmed the score and recorded it to the ODK form. The circumference measures were collected at the same marked midpoint and lymphedema stage was determined using criteria described by Dreyer, G., Addiss, D., Dreyer, P., Noroes, J. [24].

### Analysis

A one-way random effects model where people effects are random was used to determine the intra-class correlation coefficient (ICC) for the three measures and results for average measure consistency are reported. An ICC score represents the variation between repeated measures where a score of 1.0 would be a perfect score, i.e. all three measures were exactly the same (absolute correlation). Scores over 0.7 are considered good and scores over 0.9 are excellent [25]. Pearson’s Chi-Square was used to assess the ICC scores for each team over repeated instances. A coefficient of variation (CV) was calculated as the ratio of the standard deviation to the overall mean and reported as a percentage. A CV represents the dispersion of data around the mean and may be used to compare the agreement between repeated measures, a lower percentage indicating less variation (more agreement) between the measures. A lymphedema stage was allocated to each leg according the criterion described by Dreyer, G., Addiss, D., Dreyer, P., Noroes, J. [24] with stages 3 and 4 grouped together as moderate, and stages 5 – 7 as severe lymphedema. The CV was calculated for unaffected legs (stage 0), moderate and severe lymphedema. Indurometer scores were collected on both legs of each participant, with each leg serving as one ‘subject’ for the analyses. Data from the electronic tablets were uploaded each day to a secure server and later downloaded as excel files. All analyses were performed using SPSS Statistics (IBM Corp. version 24).

## Results

### Participants

In Bangladesh 146 people were included in the study but due to availability of the devices data could not be collected at baseline. In Ethiopia 126 people were included and data were available at all time points. The study was conducted over 24 weeks and at final follow-up there was a retention rate of 82.6% in Bangladesh (n = 133) and 81.8% in Ethiopia (n=103). There were no significant between country differences at baseline for factors known to affect Indurometer measures such as gender and lymphedema stage, see Table 1.

Each team measured both legs of between 51 and 79 participants at each time-point. The ICC was calculated for each team and by each leg of the participants and the results are given in Table 2. Overall the data teams achieved very good to excellent ICC scores at every instance with scores between 0.829 (95% CI 0.730, 0.896) (Ethiopia team 1 at 4-week follow-up) and 0.992 (95% CI 0.989, 0.995) (Bangladesh team 1 at 24-week follow-up). There were no significant differences in the ICC scores for each team over time (Pearson’s Chi-Square, Bangladesh p=0.199, Ethiopia p=0.213). Figure 2 plots the ICC scores for each team on a radar graph where lines closer to the rim represent more agreement between repeated measures.

In Bangladesh CV scores show that agreement between measures was highest among unaffected legs (range 16% - 22%) and the most variability was found in measures of severe lymphoedema (range 19% - 39%). In Ethiopia CV scores showed no distinct pattern of agreement for lymphedema stage (range 15% - 32%). CV scores for the data collection teams in each country can be seen in Figures 2a) Bangladesh and 2b) Ethiopia.

## Discussion

The research teams in Bangladesh and Ethiopia all achieved ICC scores over 0.8 with most instances over 0.9 and this is similar to ICC scores for the same device when used at the same measurement point in healthy young people in Myanmar and Australia (range 0.906 (0.857–0.955) - 0.942 (0.910, 0.973)) [18]. The CV scores were higher than the first published report on the same device by Pallotta, et al. [16] on BCRL lymphedema of the arm (range 7%-16%) but this study did not report on the stage of lymphedema. A further study by the same authors reported CV values of between 26% - 31% on moderate to severe arm lymphedema [17] which are similar to the CV scores reported here for moderate to severe leg lymphedema.

Three previous studies also compared reliability of the Indurometer with a mechanical tonometer and all found the mechanical device had less variability than the new digital version [13, 16-18], however the mechanical tonometer had some operational limitations and is no longer available. In particular the load was applied to the indenter using gravity, so any small deviation from holding the device upright affected the accuracy of the measure. Additionally, the result had to be read from two analogue dials with the device still in place. When using the Indurometer, the operator applies the load and a force sensor detects when the equivalent to 200-g has be attained, the measure is captured and displayed on the LED screen. This means the device can be used at any angle and the operator can remove the device from the skin to easily read the score.

There have been no large-scale studies to establish normative ranges for tissue tonometry scores and assessment of tissue compressibility is best used to compare between limb differences or within-person changes over time, rather than determining deviation from a ‘norm’. A previous study which collected tonometry measures at the anterior and posterior thigh and at mid-calf, and which compared three difference tonometers, showed that variability in repeated scores was greatest at the anterior thigh and least so at the mid-calf for all devices [18]. This may be due to the normal variation in fat layers at each site. The mid-point of the calf lies over a musculotendinous junction and generally has less fat than the thigh, so this measurement point may be less affected by changes in overall body fat over time and more sensitive to small changes in tissue composition [20]. The Indurometer was able to detected covert lymphatic dysfunction at mid-calf in young people infected with LF in Myanmar but did not detect significant differences at the thigh points [19].

In lymphedema, tissue compressibility scores should be interpreted in conjunction with other objective assessments such as lymphedema stage. In the very early fluid-rich stages, the tissue is soft and easily compressible and tonometry scores will be high. In the middles stages the gradual build-up of fibrous tissues makes the tissue feel stiff and immobile. Adipose cells may also respond by accumulating and proliferating, and this fatty induration gives the tissue a spongy quality which rebounds immediately after compression. In later stages, hyperkeratosis and sclerosis of the skin will increase tissue stiffness further [4]. Therefore, in early lymphedema when the swelling is very soft higher Indurometer scores can be expected and will decrease if the fluid is removed. In later stages when the tissue has become fibrotic, lower scores can be expected and if effective treatment softens the tissues then Indurometer scores will increase. Any assessment of tissue compressibility or stiffness should also consider known moderating factors such as age, gender and body composition [20]. Mean Indurometer scores provided here have not accounted for these factors and should not be considered to be indicative of any particular stage of lymphedema. Furthermore, because three repeat measures were used for all analyses, the Indurometer scores reported here apply only to an average of three measures. It can expected that single measures will have less repeatability and using an average of two or three measures is recommended.

One limitation to the study was the late arrival of the Indurometers in Bangladesh which were not available during baseline data collection. The data collection teams had some initial training in operation of the device but no field support once it was implemented at first follow up. In Ethiopia the device was available from the beginning of data collection and the teams had field support on hand at baseline and at the 4-week follow-up. There were also some minor issues with operation of the Indurometers and all teams reported episodes of multiple error messages or problems with recharging the battery. Nevertheless, the Indurometer is a very simple device to operate and can be used by anyone after a short period of training and practice. It has demonstrated good to excellent reliability when used to measure upper-limb lymphedema, the lower-limbs of healthy young populations, and now, in an adult population with overt LF- or podoconiosis-related leg lymphedema. The Indurometer is a useful device to assess tissue compressibility in lymphedema and is suitable for use in clinical research in multiple settings.

## Conclusions

The complex nature of lymphedema precludes a single defining diagnostic tool or feature to determine lymphedema status. A combination of methods is ideal in developing a full picture of all connective tissue changes and tissue tonometry has an important place in this assessment tool-kit. The Indurometer is a field- and user-friendly device which has proved reliable in several populations.

## Acknowledgments

Grateful thanks are given to the Bangladesh MOHFW and the Ethiopia FMOH, CIPRB and NaPAN for their support in the design, development and implantation of the trial. We also sincerely thank our research assistants in each country. In Bangladesh, Muhammad Mujibur Rahman, Abdullah Al Kawsar, Mahmud Hasan, Islam Md. Faizur Rashid, Farzana Mohsin, Imran Nazir, Mst Mansia Khatun, S. M. Raihan Nesar Jewel, Dulal Hossain, Kazi Burhan Uddin, Samid Siddiqi. In Ethiopia, Mossie Tamiru Wondimeneh, Getnet Agidew, Engdaw Getachew Abera, Gizachew Mezgeb Asfaw, Shimels Bahiru Adugna, Wondie Ayele Golnetie, Aklilu Ayehu Tassew, Fentaw Ayalew Nebebe, Kassahun Ewnetu Melese, Ewunetyhun Wale Gebeyehu, Ambachew Birhan Adugna.

### Funding

The study was funded by the United Kingdom Department for International Development (DFID) and Glaxo Smith Kline (GSK). Neither funding body had any input into study design, data collection, management, analysis or interpretation of data.

### Author Disclosure Statement

No competing financial interests exist.

## References

1. International Society of Lymphology, *The diagnosis and treatment of peripheral lymphedema: 2016 consensus document of the international society of lymphology.* Lymphology, 2016. **49**(4): p. 170-184.

2. Rockson, S.G. and K.K. Rivera, *Estimating the population burden of lymphedema*. 2008. p. 147-154.

3. Deribe, K., Cano, J., Trueba, M.L., Newport, M.J., Davey, G., *Global epidemiology of podoconiosis: A systematic review.* PLOS Neglected Tropical Diseases, 2018. **12**(3): p. e0006324.

4. Lawenda, B.D., Mondry, T.E., Johnstone, P.A.S., *Lymphedema: A primer on the identification and management of a chronic condition in oncologic treatment.* CA Cancer Journal for Clinicians, 2009. **59**(1): p. 8-24.

5. Douglass, J., Kelly-Hope, L., *Comparison of Staging Systems to Assess Lymphedema Caused by Cancer Therapies, Lymphatic Filariasis, and Podoconiosis.* Lymphat Res Biol, 2019.

6. Douglass, J., Graves, P., Gordon, S., *Self-Care for Management of Secondary Lymphedema: A Systematic Review.* PLOS Neglected Tropical Diseases, 2016. **10**(6): p. e0004740.

7. Clodius, L., Deak, L., Piller, N.B., *A new instrument for the evaluation fo tissue tonicity in lymphoedema.* Lymphology, 1976. **9**(1): p. 1-5.

8. Kar, S.K., Kar, P.K., Mania, J., *Tissue tonometry: a useful tool for assessing filarial lymphedema.* Lymphology, 1992. **25**(2): p. 55-61.

9. Moseley, A.L., Piller, N.B., Carati, C.J., *The effect of gentle arm exercise and deep breathing on secondary arm lymphedema.* Lymphology, 2005. **38**(3): p. 136-45.

10. Bagheri, S., Ohlin, K., Olsson, G., Brorson, H., *Tissue tonometry before and after liposuction of arm lymphedema following breast cancer.* Lymphatic Research and Biology, 2005. **3**(2): p. 66-80.

11. Lau, R.W., Cheing, G.L., *Managing postmastectomy lymphedema with low-level laser therapy.* Photomedicine and Laser Surgery, 2009. **27**(5): p. 763-9.

12. Douglass, J., Immink, M., Piller, N., Ullah, S., *Yoga for women with breast cancer-related lymphoedema: A preliminary 6-month study.* Journal of Lymphoedema, 2012. **7**(2): p. 30-38.

13. Moseley, A.L., Piller, N.B., *Reliability of bioimpedance spectroscopy and tonometry after breast conserving cancer treatment.* Lymphatic Research and Biology, 2008. **6**(2): p. 85-87.

14. van Zanten, M., Kean, B., *Use of a two-layer compression system in severe bilateral leg lymphoedema with ulceration: A case report.* Journal of Lymphoedema, 2013. **8**(2): p. 25.

15. Gordon, S., Melrose, W., Warner, J., Buttner, P., Ward, L., *Lymphatic filariasis: a method to identify subclinical lower limb change in PNG adolescents.* PLOS Neglected Tropical Diseases [electronic resource], 2011. **5**(7): p. e1242.

16. Pallotta, O., McEwen, M., Tilley, S., Wonders, T., Waters, M., Piller, N., *A new way to assess superficial changes to lymphoedema.* Journal of Lymphoedema, 2011. **6**(2).

17. Vanderstelt, S., Pallotta, O.J., McEwen, M., Ullah, S., Burrow, L.., Piller, N., *Indurometer vs. Tonometer: Is the Indurometer Currently Able to Replace and Improve Upon the Tonometer?* Lymphat Res Biol, 2015. **13**(2): p. 131-6.

18. Douglass, J., Graves, P., Gordon, S., *Intrarater Reliability of Tonometry and Bioimpedance Spectroscopy to Measure Tissue Compressibility and Extracellular Fluid in the Legs of Healthy Young People in Australia and Myanmar.* Lymphat Res Biol, 2017. **15**(1): p. 57-63.

19. Douglass, J., Graves, P., Lindsay, D., Becker, L.., Roineau, M., Masson, J., Aye, N.N., Win, S.S., Wai, T., Win, Y.Y., Gordon, S.J., *Lymphatic Filariasis Increases Tissue Compressibility and Extracellular Fluid in Lower Limbs of Asymptomatic Young People in Central Myanmar.* Tropical Medicine and Infectious Disease, 2017. **2**(4): p. 50.

20. Douglass, J., Graves, P., Gordon, S., *Moderating Factors in Tissue Tonometry and Bioimpedance Spectroscopy Measures in the Lower Extremity of Healthy Young People in Australia and Myanmar.* Lymphat Res Biol, 2018. **16**(3): p. 309-316.

21. Hartung, C., Lerer, A., Anokwa, Y., Tseng, C., Brunette, W., Borriello, G., *Open data kit: tools to build information services for developing regions*, in *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development*. 2010, ACM: London, United Kingdom. p. 1-12.

22. Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B. J., Hrobjartsson, A., Roberts, C., Shoukri, M., Streiner, D.L., *Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed.* J Clin Epidemiol, 2011. **64**(1): p. 96-106.

23. World Medical Association, *World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects.* Jama, 2013. **310**(20): p. 2191-4.

24. Dreyer, G., Addiss, D., Dreyer, P., Noroes, J., *Basic Lymphoedema Management, Treatment and Prevention Problems Associated with Lymphatic Filariasis*. 2002, USA: Hollis Publishing Company.

25. Fleiss, J.L., *Reliability of Measurement*, in *The Design and Analysis of Clinical Experiments*. 1999, John Wiley & Sons, Inc. p. 1-32.

## Figure legends

#### Figure 1. a) the Indurometer, with b) an arrow showing the direction of application. The measure is captured when an equivalent of 200g has been applied.

*Figure 2: Intra-class correlation co-efficient (ICC) scores for both teams at all time points. Lines closer to the rim indicate a higher ICC, ie more reliability.*

*Figure 3a) Coefficient of Variation (CV) for data collection teams in Bangladesh by lymphoedema stage.*

*Figure 2b) Coefficient of Variation (CV) for data collection teams in Ethiopia by lymphoedema stage.*

#### Supplementary Figure 1 a). Operation of the Indurometer

#### Supplementary Figure 1 b). Daily recalibration of the Indurometer