aranzmedical

Electronic Wound Measurement Systems
A review of, and template for, the development and validation

by CM Frampton PhD, MA Nixon MD & RW Watson PhD

Abstract

A systematic review of the research undertaken to determine the statistical attributes (including accuracy, precision and longitudinal accuracy and precision) of electronic wound assessment systems has been conducted. The review shows that by using wound models in research, random and systematic errors that occur as part of clinical use are largely ignored and optimistic unrealistic conclusions are drawn. Further, studies often ignore users, the nonspecific nature of the wound boundary, patient characteristics, wound types and characteristics and the variation associated with the temporal measurement of wound healing as potential sources of measurement variation. Much of the research also uses correlation analysis where measurements from non-electronic wound measurement techniques, that are known to be imprecise and inaccurate, are compared to those of the electronic system. This paper provides guidance on how to recognize these methodological shortcomings and how to develop a sequential process for the statistical validation of a wound assessment device.

Introduction

In recent years there has been a prolific development of devices and applications to allow the photographing, measurement and documentation of wounds[1,2,3]. These devices have been developed to address firstly the clinical and legal concerns around poor quality documentation and inconsistent measurement, and secondly to improve evidence based treatment strategies and to drive continuous improvement in wound-care organisations[4,5,6]. Some of these devices have application in clinical research. Traditional methods including ruler based length, width, and area measurements (based on length and width measurements) have been found to lack both accuracy and precision, whereas acetate tracing lacks both user and patient acceptability[1,2,7].

Electronic Wound Measurement (EWM) devices were developed to measure wounds accurately and precisely, and to make the wound assessment process easier and more streamlined. However, the quality of the research underpinning the effectiveness of these systems is highly variable with few systems going through a systematic program of research. This situation may be less concerning for institutions and clinicians wanting to see an improvement over traditional techniques, as any electronic technique is likely to result in improvement, but it is concerning to those seeking to make evidence based decisions on patient treatment or organizational improvement.

This poster reviews the development and validation research that has been performed on EWM devices in recent times, highlighting deficits in the extent and quality of studies. It proposes a template for conducting this research.

Background

Area reduction is the primary measure most commonly used in the clinical setting to gauge healing progress[1,2,4,5]. Area measurement is common amongst EWM devices and is therefore, the focus of this paper, even though the physiology of wound-healing suggests volume and depth may also be useful markers[1,2].

Key requirements of an EWM device are the:
- accuracy of single measurements
- accuracy of longitudinal measurements
- precision of single measurements
- precision of longitudinal measurements
- smallest clinically relevant change that can be detected
- user acceptability
- patient acceptability

Accuracy:
- The extent to which a measurement device is nonbiased and consistently measures the true area – see Figure 1
- **Longitudinal accuracy** is the extent to which the device measures the true absolute or percentage change over time, this assumes consistent accuracy over wound size – see Figures 2 and 3
- Inaccuracy is often referred to as bias
- Potential sources of inaccuracy (or bias) derive from sources that include the environment, device, user, patient and wound

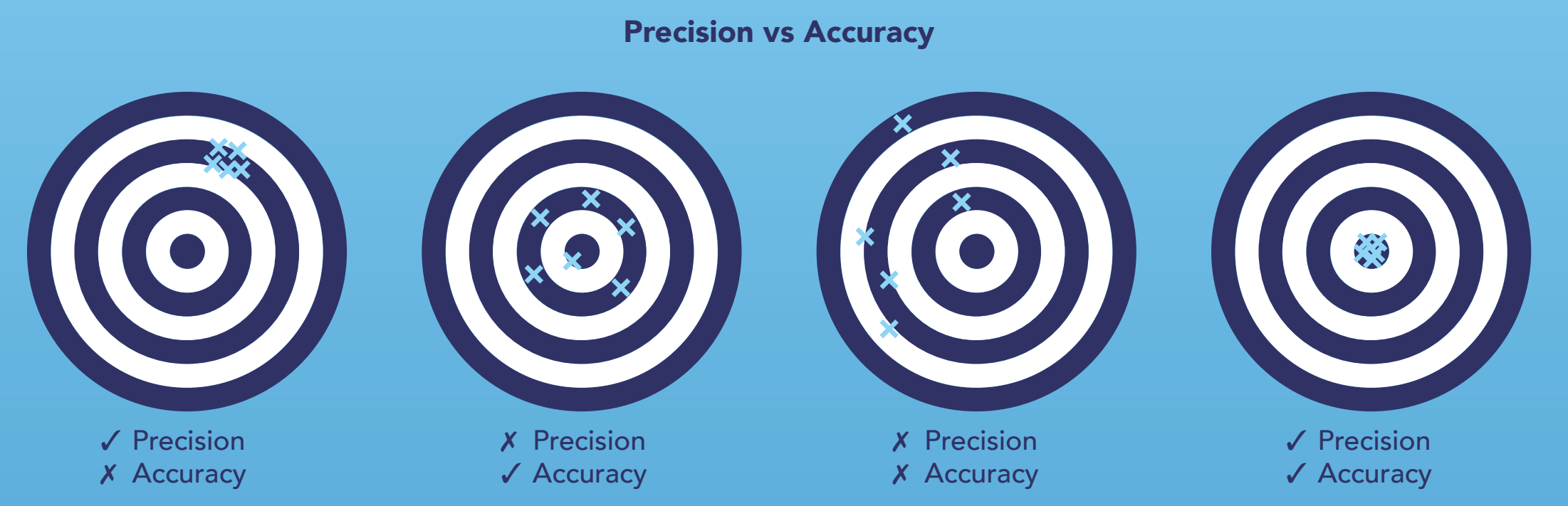

Figure 1: A graphical representation of precision and accuracy for six measurements of a wound.

Precision:
- The repeatability of the measurement device, so that it returns a consistent measure, when the wound has not changed – see Figure 1
- **Longitudinal precision** is the extent to which the measurement device provides a consistent measure of absolute or percentage change – see Figures 2 and 3
- Imprecision is often referred to as noise
- Potential sources of imprecision (or noise) derive from sources that include the environment, device, user, patient and wound

Figure 2: A graphical representation of precision and accuracy for six measurements of a wound (three at t_0 and three at t_1).

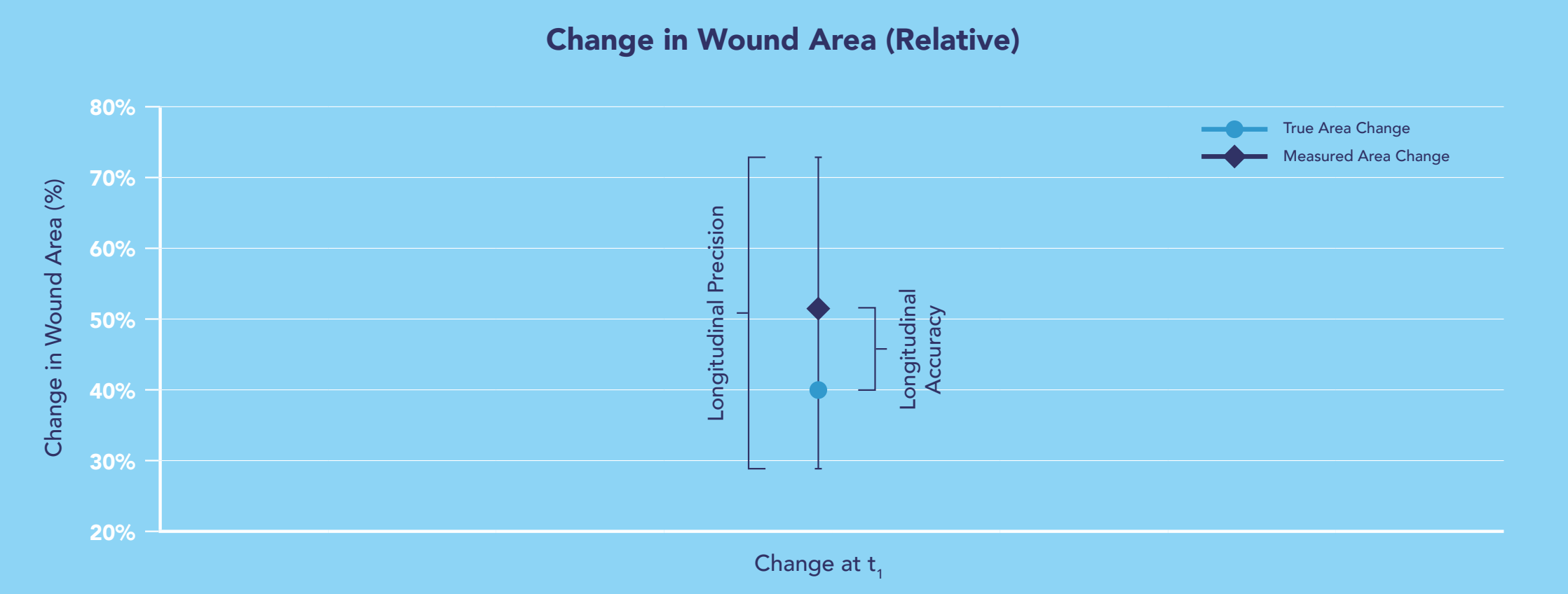

Figure 3: A graphical representation of longitudinal accuracy and precision derived from the measurements in Figure 2.

The various algorithms, imaging technologies and measurement technologies used in the devices will be associated with their own error profiles, and also external sources of variation including environment, device, user, patient and wound. These sources of variation will affect the accuracy and precision of the EWM devices to varying degrees, and examples of these variation sources include:

Source of variation	Example
Environment	Lighting
Device	Algorithms, underlying imaging and measurement technologies
User	Human factors leading to inter and intra-rater variation
Patient	Skin pigment, amount of adipose tissue and ability of the patient to cooperate
Wound	Complexity of the wound structure, etiology, depth, healing stage, wound location and moisture

Minimum clinical difference:
- The minimum change in wound area that clinicians determine as being clinically relevant in terms of the status of the wound healing
- This defines a quantum that the EWM device must be able to resolve and that is distinguishable from measurement error

User Acceptability:
- Extent to which the measurement device is easy to use, the amount of training required, suitability for the task, compatibility with existing workflows, amount of time required to perform its task, cost and ability to be implemented

Patient acceptability:
- Is the measurement device associated with pain or discomfort, for example

Template

As with other areas of medical efficacy research, in order to build the evidence pyramid for the various EWM devices, it is necessary to conduct a program of sequential studies. This program of studies should assess the ideal performance of the systems, for example with wound models in a controlled environment, gradually introduce and characterize the impact of various potential sources of variation, until the study setting clearly reflects the real world clinical scenario. A key aspect of the study program is that it must determine the impact of the potential sources of variation on longitudinal precision and accuracy.

A suitable study program should include the following sequence of studies:
- Computer simulation of core algorithms with computer based wound models
- Physical wound models in a single, controlled environment
- Physical wound models with known sources of variation, for example environmental (lighting), user (inter and intra-rater variation) sources
- Real wounds in a controlled environment
- Real wounds with known and managed sources of variation
- Real wounds longitudinal study
- Real world validation studies

Statistical Considerations

Statistical considerations for development and validation studies of an EWM device.

Accuracy is measured as the average difference between measured and the actual areas. These are often portrayed using Bland-Altman plots which demonstrate the average difference (bias) and the 95% Limits of Agreement[8]. The limits of agreement represent the variation in the differences between the measured and actual areas. An example is presented in Figure 4.

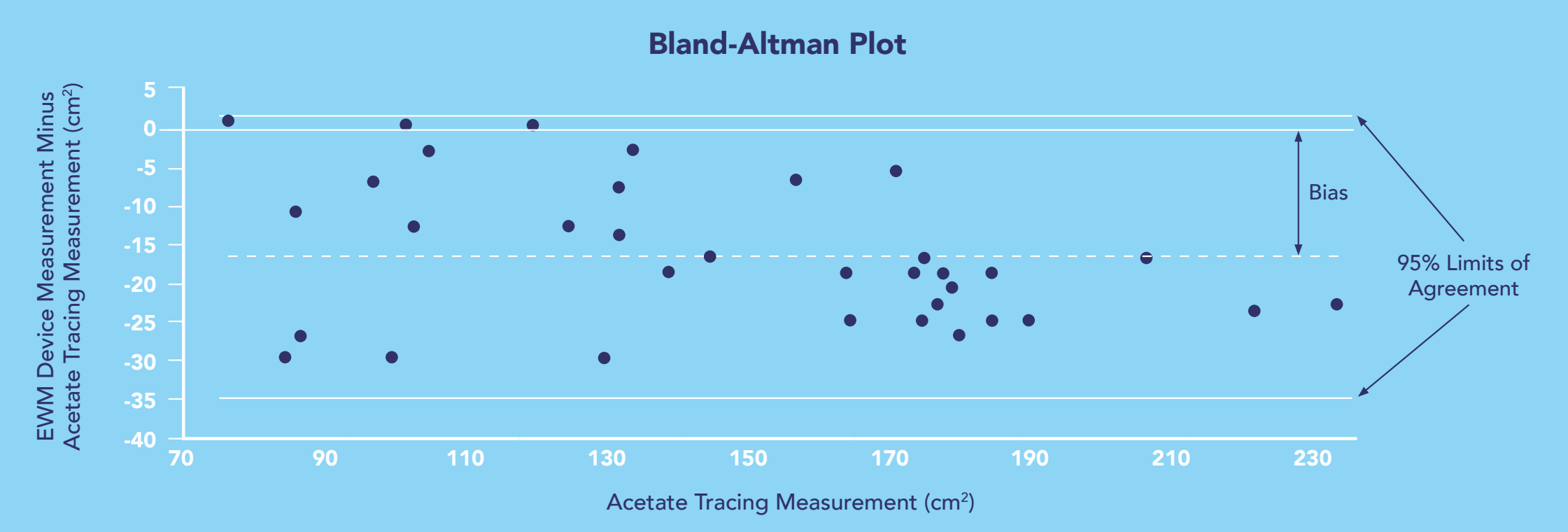

Figure 4. Bland-Altman plot showing the differences in measurement (EWM device measurement for a wound less the acetate tracing measurement for the same wound) plotted against the acetate tracing measurement. The EWM device consistently underestimates the acetate tracing measurement by 17.0 mm².

Precision is measured as the variation around a measured mean value for a single wound and is therefore independent of accuracy. Precision quantifies the repeatability of a measurement and is summarized as the standard deviation (SD) of repeat measurements of the same wound. SD is often calculated from repeat measurements from a number of different wounds. Precision is further summarized using the following statistical measures.

Coefficient of variation (CV%) represents the standard deviation of the repeat measurements (SD) adjusted for the mean of all measurements. This is calculated as SD/mean and expressed as a percentage. Larger values correspond to poorer precision.

Intra-class correlation coefficient (ICC). This summary coefficient represents the magnitude of sources of variation that have been defined and studied. In the context of wound measurement, inter and intra-rater variation is most commonly studied. The ICC is a number that varies between 0 and 1.0, or expressed as a percentage from 0 to 100%. With a perfect measuring device, all the variation in measurements is simply due to differences in individual wound areas, and the ICC =1.0, i.e. there is no inter and intra-rater variation. The ICC is calculated as the percentage of the total variation between all area measurements that is attributable to the variation between the wound areas. The method of calculation means that the interpretation of the ICC needs to be considered in the context of the actual differences in wound areas.

Minimum clinical difference (MCD). This summary statistic utilizes the measured precision of a device to calculate the minimum change in wound area that would represent a genuine change rather than representing a random difference as a consequence of poor precision. The MCD can be reduced by taking repeat measurements of the wound area on each measurement occasion. Ideally the MCD should be calculated using the precision of the device to measure change rather than the precision estimated at a single time point.

Variance component decomposition. These are standard statistical techniques which use the data from an appropriately designed study to estimate the contribution of different sources of variation (e.g. inter and intra-rater). If a study has a number of users each repeatedly measuring the area of a range of real wounds, then the CV% and the ICCs attributable to inter, intra-rater variation and both sources combined can be calculated.

Review

Method

To identify published literature studying EWM devices the authors:
1. Searched data-bases publically visible on the internet for documents describing studies:
   a. On EWM devices in general
   b. Reviewing studies on EWM devices
   c. On the devices of known vendors of EWM devices
2. Searched the websites of EWM device vendors for studies describing the effectiveness of their systems
3. Used the references in the literature in step one and two to identify further studies
4. Used the knowledge gained in the above steps to search for further publications by the principal investigators or inventors of EWM devices

While extensive efforts were made to identify all relevant literature, it is possible that individual published and unpublished studies have not been identified, but we believe the literature reviewed is representative of the research undertaken for EWM devices.

To be considered relevant as part of this review, the studies needed to describe system(s) that both imaged and measured various dimensions of the wound, and not those that simply performed electronic documentation or the recording of manual measurements.

After reviewing the studies the authors categorized each study according to the type(s) of EWM device(s) investigated and the stage in the study program template described above.

Results

Twenty six publications describing studies validating EWM devices were found.

It was found the EWM devices described in the studies could be broadly classified, based on their measurement technology, into different groups, as has been previously done by others[9,10,11]. The authors found the devices could be classified into four broad groups. These include laser assisted wound measurement devices (LAWM) using structured light to measure the surface of the wound bed, applications using commercially available 3D structure sensors (3DWM), applications using a target placed next to the wound and digital planimetry to determine the scale of the wound (2DPA), and devices using multiple point light sources with known geometry to determine the range between the wound and the measuring device (2DAR).

The following table provides an overview of the extent to which the study program has been completed on each category of device. The quality of the studies within the program is described later.

Stage in study program	LAWM	3DWM	2DPA1	2DPA2	2DAR1	2DAR2
Computer simulation – core algorithms		✓[12]✓[13]✓[14]	✓[15]			
Wound models – controlled environment	✓[9]✓[10]✓[16]	✓[10]			✓[17]✓[18]	✓[19]
Wound models – known sources of variation	✓[10]✓[20]✓[32]	✓[10]				
Real wounds - controlled environment	✓[21]✓[22]	✓[11]			✓[18]	
Real wounds - known sources of variation						
Real wounds longitudinal study	✓[6]✓[23]✓[24]					
Real world validation studies	✓[25]✓[26]✓[27] ✓[28]✓[29]✓[30] ✓[31]					

Discussion

The above table demonstrates that:
- Almost all devices have had at least one study conducted in an effort to quantify attributes of the device.
- Almost all devices have conducted at least one early stage wound model study, but few have been through a program that includes later stage real wound studies.
- Few devices have had studies conducted where known sources of variation were explicitly evaluated. When these were conducted they usually only explored the inter and intra-rater sources of variation.
- For almost all devices the attributes were not evaluated over time.

The reviewed studies cover a spectrum of designs and objectives, however some themes that were observed were:
- All have as their primary objective the determination of the statistical attributes of the device(s).
- Only a minority of studies also addressed some aspects of the user acceptability.
- A key distinction was whether early stage wound model studies or later stage real wound studies were conducted. Early stage studies are important for verification and refinement of the measuring device, however the later stage testing is required to validate the real world performance and acceptability.
- Where only early stage wound model studies were conducted there was often no discussion of the impacts of the different sources of real world variation on the performance of the device.
- Many studies compared their EWM device to another EWM device or a traditional method, and often this choice of comparator was inappropriate. For example, some studies compare the EWM device to wound areas derived from length x width measurements. While such traditional methods may be the standard of care, they do not serve as an appropriate comparator. A clear definition and rationale for the choice of comparator method should be provided.
- Many studies on real wounds limited themselves to a particular wound type, e.g. diabetic foot ulcers, and implied there is generalizability from these results to other wound types. If a device is to have broad clinical application for the measurement of wounds then a range of wound types located at various anatomical sites needs to be part of the evaluation.

The reviewed studies utilized various forms of statistical analysis and it was observed:
- Precision and accuracy are often confused and misunderstood.
- Some erroneously implied that the level of accuracy and precision at a single point in time is sufficient to derive the longitudinal accuracy and precision over time.
- Inter and intra-rater results are frequently addressed. For the device to be useful these need to be undertaken on real wounds of an appropriate type, using clinicians who will ultimately use the device. Earlier stage testing of inter and intra-rater variation on models is useful but is likely to underestimate the true and important variation in a clinical environment.
- ICCs are occasionally used to quantify inter and intra-rater variation. It is critical that the ICCs are interpreted in the context of the wound areas used in the study. For example, if you are testing on wound models which are all of similar size this will underestimate the ICC. In many circumstances the coefficient of variation is a better summary of the inter and intra-rater variation.
- Correlation coefficients are sometimes used to compare devices and methods. These are inappropriately used to summarise the agreement. For example, two devices whose measurements differ by a fixed amount will have a correlation coefficient of 1.0 but will not have perfect agreement.
- Only one study considered and determined the minimum clinical difference[24].

Conclusions

In almost all cases no structured approach has been employed in the development and validation research to assess the performance of EWM devices. However, the LAWM device and to a lesser degree the 3DWM device have been subject to many of the elements of a suitable sequential study program, as described in this poster. While some of the individual studies reviewed did have statistical limitations some did however, identify important individual components appropriate for the assessment of the statistical attributes and address these components in a robust and statistically valid manner. Amongst the studies reviewed there were inconsistent approaches to the design, analysis and interpretation of validation studies. There is a need for the development of a structured, statistically robust framework for the development of EWM devices. Our poster presents a platform for such a framework.

References

1. *Diagnostic and Prognostic Utility of Non-Invasive Multimodal Imaging in Chronic Wound Monitoring: a Systematic Review*, R. Mukherjee, S. Tewary & A. Routray, J Med Syst. 2017, 41(3)
2. *Methods to assess area and volume of wounds – a systematic review*, L. B. Jørgensen, J. A. Sørensen, G. B. E. Jemec, and K. B. Yderstræde, Int. Wound J. 2016, 13(4): 540–553
3. *The Evolving Field of Wound Measurement Techniques: A Literature Review*, R. Khoo & S. Jansen, Wounds 2016; 28(6): 175-181
4. *Wound measurement: can it help us to monitor progression to healing?* M. Flanagan, J Wound Care 2003, 12(5): 189-94
5. *A multicenter study of percentage change in venous leg ulcer area as a prognostic index of healing at 24 weeks*, J. Kantor & D. J. Margolis, Br J Dermatol 2000, 142(5): 960-4
6. *Technological advances in wound bed measurements*, M. Romanelli, Wounds 2002, 14(2): 58–66
7. *Digital planimetry results in more accurate wound measurements: a comparison to standard ruler measurements*, L. C. Rogers, N. J. Bevilacqua, D. G. Armstrong & G. Andros, J Diabetes Sci Technol 2010, 4:799–802.
8. Statistical methods for assessing agreement between two methods of clinical measurement, J.M. Bland & D.G. Altman, Lancet. 327(8476): 307–10.
9. *Validation of a laser-assisted wound measurement device in wound healing model*, R. S Constantine, J. D. Bills, L. A. Lavery & K. E. Davis, Int Wound J 2014, doi: 10.1111/iwj.12328.
10. *Pilot Study to evaluate a novel 3-D wound measurement device*, J. D. Bills, S. J. Berriman, D. L. Noble, L. A. Lavery & K. E. Davis, Int. Wound J. 2015, 13(6), 1372-1377
11. *The reliability of a Novel Mobile 3-Dimensional Wound Measurement Device*, E.L Anghel, A. Kumar, T. E. Bigham, K. M. Maselli, J. S. Steinberg, K. K. Evans, & P. J. Kim 2016, Wounds Epub, 28, Aug.
12. *Method and System for Wound Assessment and Management*, K. Wu, ACS Clinical Congress 2014
13. *Mobile Structure Sensor for Real-Time 3D Wound Assessment : Ex-Vivo Validation using Wound Phantoms*, K. Wu, WOCN Annual Conference 2015
14. *Feasibility Of 3D Structure Sensing For Accurate Assessment Of Chronic Wound Dimensions*, K. L. Wu, K. Maselli, A. Howell, D. Gerber, E. Wilson, P. Cheng, P. Cw. Kim & O. Guler, Computer Assisted Radiology and Surgery International Congress 2015
15. *Design of a Smartphone Application for Automated Wound Measurements in Home Care*, J. Budman, K. Keenahan, S. Acharya, & G. A. Brat , Connected Health Symposium 2015, www.iproc.org/2015/1/e16
16. *Assessment of accuracy and repeatability on wound models of a new hand-held, electronic wound measurement device*, M. Nixon, T. Rivett & B. Robinson, World Union of Wound Healing Societies Congress (WUWHS) 2012, Yokohama, Japan.
17. *Applying 21st Century Imaging Technology to Wound Healing: An Avant-Gardist Approach*, F. L. Bowling, J. Paterson , A. Ndip, J Diabetes Sci Technol. 2013 7(5):1190-4.
18. *Clinical usability of a wound measurement device*, M. Nemeth, S. Sprigle, A. Gajjala. 36th Annual American Spinal Injury Conference, Nashville, TN, 2010.
19. *Accuracy and Reproducibility of the Wound Shape Measuring and Monitoring System*, D. Langemo, J. Spahn, L. Snodgrass, Adv Skin Wound Care. 2015, 28(7):317-23.
20. *Simulated Wound Assessment using digital planimetry vs three-dimensional cameras – implications for clinical assessment*, K. J. Williams, V. Sounderajah, B. Dharmarajah, A. Thapar, & A. H. Davies.Ann Vasc Surg. 2017 Feb.
21. *The Reliability of a Handheld Wound Measurement and Documentation Device in Clinical Practice*, C. E. Hammond & M. A. Nixon, J Wound Ostomy Continence Nurs., 38(3):260-4
22. *Measurement of Diabetic Foot Ulcers using SilhouetteMobile*, D. Nixon, A. Philps , L. Jonsson & M. Nixon, Journal of Wound Technology, 2009, 3, pp 37-38
23. *Reliability and Efficiency of Novice Raters Performing Surface Area Measurement with Two Different Methods on Simulated Wound Models*, K. Albaugh, Clinical Electrophysiology & Wound Management 2014, 28(1)
24. *Intra- and Inter-rater Reliability of Wound Surface Area Measurement Over Time with Photo Digital Planimetry*, K. Albaugh, M. Heim, J. Martin, P. Mechling, K. Neff, & J. Peters, APTA, Feb 2017,
25. *Spray-applied cell therapy with human allogeneic fibroblasts and keratinocytes for the treatment of chronic venous leg ulcers:a phase 2, multicentre, double-blind, randomised, placebocontrolled trial*, R. S. Kirsner, W. A. Marston, R. J. Snyder, T. D. Lee, D. I. Cargill, H. B. Slade, The Lancet, 2012, 380(9846), pp 977-985.
26. *Wound initial size reduction as a predictor of full wound healing in patients with leg ulcers*, M. Wikkerink, & G. Koel , WCS Nieuws, 2011, 27(4), pp 34-36.
27. *Effect of TNP on the microbiology of venous leg ulcers: a pilot study*, M. Khashram, P. Huggan, R. Ikram, S. Chambers, J. Roake, & D. Lewis, Journal of Wound Care, April 2009, 18(4), pp 184-187.
28. *Cellular versus acellular matrix devices in treatment of diabetic foot ulcers: study protocol for a comparative efficacy randomized controlled trial.* H. Lev-Tov, L. Chin-Shang, S. Dahle, & R. R. Isseroff, Trials 2013 14:8.
29. *Uncultured adipose-derived regenerative cells (ADRCs) seeded in collagen scaffold improves dermal regeneration, enhancing early vascularization and structural organization following thermal burns*, P. Foubert, S. Barillas, A. D. Gonzalez, Z. Alfonso, S. Zhao, I. Hakim I, C. Meschter, M. Tenenhaus, & J. K. Fraser, Burns 2015, 41(7):1504-16
30. *Development of a combined radiation and full thickness burn injury minipig model to study the effects of uncultured andipose-derived regenerative cell therapy in wound healing*, P. Foubert, M. Doyle-Eisele, A. Gonzalez, F. Berard, W. Weber, D. Zafra, Z. Alfonso, S. Zhao, M. Tenenhaus, & J. K. Fraser Int J Radiat Biol 2016, 93(3), 340-350.
31. *Healing Rates in a Multicenter Assessment of a Sterile, Room Temperature, Acellular Dermal Matrix Versus Conventional Care Wound Management and an Active Comparator in the Treatment of Full-Thickness Diabetic Foot Ulcers*, J. Walters, D. S. Cazzell, H. Pham, D. Vayser, and A. Reyzelman, Eplasty. 2016, 16:e10.
32. *The Reliability of a Novel Mobile 3-dimensional Wound Measurement Device*, K. E. Davis, F. C. Constantine, E. C. MacAslan, J. D. Bills, D. L. Noble, and L. A. Lavery, J. of Diabetes Sci, Technol. 7(5), 1161-1166