electronics

Article

Evaluation of Two Digital Wound Area Measurement Methods Using a Non-Randomized, Single-Center, Controlled Clinical Trial

Lorena Casanova-Lozano [1,*,†], David Reifs-Jiménez [1,*,†], Maria del Mar Martí-Ejarque [2], Ramon Reig-Bolaño [1] and Sergi Grau-Carrión [1]

1 Digital Care Research Group, Centre for Health and Social Care Research (CESS), Universitat de Vic–Universitat Central de Catalunya, 08500 Vic, Spain; ramon.reig@uvic.cat (R.R.-B.); sergi.grau@uvic.cat (S.G.-C.)
2 Grupo Quirónsalud, Hospital Universitari Sagrat Cor, 08029 Barcelona, Spain; mmmarti@quironsalud.es
* Correspondence: lorena.casanova@uvic.cat (L.C.-L.); david.reifs@uvic.cat (D.R.-J.)
† These authors contributed equally to this work.

Abstract: A prospective, single-center, non-randomized, pre-marketing clinical investigation was conducted with a single group of subjects to collect skin lesion images. These images were subsequently utilized to compare the results obtained from a traditional method of wound size measurement with two novel methods developed using Machine Learning (ML) approaches. Both proposed methods automatically calculate the wound area from an image. One method employs a two-dimensional system with the assistance of an external calibrator, while the other utilizes an Augmented Reality (AR) system, eliminating the need for a physical calibration object. To validate the correlation between these methods, a gold standard measurement with digital planimetry was employed. A total of 67 wound images were obtained from 41 patients between 22 November 2022 and 10 February 2023. The conducted pre-marketing clinical investigation demonstrated that the ML algorithms are safe for both the intended user and the intended target population. They exhibit a high correlation with the gold standard method and are more accurate than traditional methods. Additionally, they meet the manufacturer's expected use. The study validated the performance, safety, and usability of the implemented methods as a valuable tool in the measurement of skin lesions.

Keywords: wounds and injuries; body surface area; machine learning

Citation: Casanova-Lozano, L.; Reifs-Jiménez, D.; Martí-Ejarque, M.d.M.; Reig-Bolaño, R.; Grau-Carrión, S. Evaluation of Two Digital Wound Area Measurement Methods Using a Non-Randomized, Single-Center, Controlled Clinical Trial. *Electronics* **2024**, *13*, 2390. https://doi.org/10.3390/electronics13122390

Academic Editor: Sergey Y. Yurish

Received: 11 May 2024
Revised: 13 June 2024
Accepted: 14 June 2024
Published: 18 June 2024

1. Introduction

A wound occurs when the skin tissue undergoes damage or breaks down. This initiates a process of regeneration known as cicatrization, which can range from hours to years or may not occur at all. A wound is deemed chronic when the repair process is significantly prolonged or lacks orderly progression. Conversely, acute wounds typically heal gradually, in accordance with the size and type of wound, usually within a short timeframe [1]. Among the most commonly treated lesions are vascular ulcers (both venous and arterial), diabetic foot ulcers, and pressure ulcers [2].

In Europe, an estimated 1.5 to 2 million people suffer from acute or chronic wounds. These types of skin injuries are typically managed in hospitals or community settings, such as primary health centers or private homes with the assistance of visiting nurses. Wounds present a pressing clinical challenge due to their significant impact on both patients and the healthcare system. Patients experience a marked decrease in quality of life as a result of the physical, cognitive, and social effects of skin injuries and their treatment [3]. Moreover, skin injuries have a substantial impact on healthcare costs due to their high prevalence, recurrence, and diversity [4]; the time invested by nurses and other healthcare professionals; and the healthcare expenses incurred from frequent dressing changes and potential wound complications [3]. Furthermore, influenced by factors such as population aging, diabetes, obesity, and bacterial resistance to antibiotics (resulting in persistent infections), wounds are expected to remain a significant clinical, social, and economic concern in the years ahead [5].

1.1. Clinical Significance of Calculating the Area of Skin Lesions

Measurement of wound size is integral to managing individuals with skin lesions, spanning from the initial assessment and classification of the wound to the selection of therapeutic strategies and evaluation of wound progression [6]. According to Gethin, G. and Nicholas, E. [7,8], skin injury measurement holds clinical significance in understanding the healing status. Moreover, measuring the surface area of a skin lesion is recommended in national and international clinical guidelines for wound care and management [9,10].

The framework for wound assessment in clinical practice is the "Wound Assessment Triangle", endorsed by the World Union of Wound Healing Societies [11]. This triangle focuses on three aspects: the wound bed, the wound edge, and the perilesional skin, with wound size measurement being a component of the wound bed assessment. Consequently, wound size is among the parameters that should be documented during wound assessment, whether utilizing the "Wound Assessment Triangle" framework or other validated wound evolution assessment scales, such as the PUSH 3.0 (Pressure Ulcer Scale for Healing) or the RESVECH 2.0 (Expected Results of the Assessment and Evolution of Chronic Wound Healing Index 2.0) [12].

Therefore, achieving a standardized, accurate, and straightforward measurement of wounds is crucial for the optimal management of individuals with this pathology.

1.2. Previous Work

To determine area measurement, some solutions are already available on the market. Some of these are based on manual methodologies, such as Visitrak™ (Smith & Nephew, Watford, UK), a tool that involves tracing the wound's outline on a transparent sheet to obtain its measurement. This process is typically time-consuming, inconvenient for the patient, and often less accurate.

Other methods utilize more advanced technologies, such as 3D imaging [13], found in commercial products like SilhouetteMobile (Aranz Medical Ltd., Christchurch, New Zealand), WoundZoom (WoundZoom, Inc., Stevens Point, WI, USA), InSight (eKare, Fairfax, VA, USA), or WoundVue [14], which employs infrared technology. These solutions often require external hardware that must be integrated, increasing their cost.

Additionally, several studies have evaluated the efficacy of some existing techniques and products. Kuang B. et al. [15] used 150 images of diabetic foot ulcers to calculate the area, depth, and volume with a mobile application called NDKare, comparing its accuracy with Visitrak and WoundVue. NDKare demonstrated a very high correlation with the other methods. Foltynski, P. et al. [16], compared the AreaMe software with Visitrak and SilhouetteMobile in 108 wounds measured five times with each device. The results showed that AreaMe was more accurate than Visitrak but less accurate than SilhouetteMobile.

Considering that time is also an important factor when measuring a wound, Mohammed H.T. et al. [17] demonstrated that the response time in 91 patients and 115 wounds using Swift Medical Inc., a mobile application for wound area measurement, was shorter than with traditional methods.

Previous solutions have demonstrated area calculation using economically costly tools or those requiring patient contact. This study introduces a cost-effective alternative, validating area measurement using a calibration system with a biocompatible marker suitable for clinical use and an AR system capable of measuring the area without patient contact. By leveraging Apple's ARKit [18], a cutting-edge tool in the field of AR, we provide a powerful platform that enables the manipulation and adaptation of existing code to meet specific functional requirements. Utilizing the sophisticated camera technology integrated into iPhone and iPad devices, ARKit demonstrates its ability to detect and track various planes within the real physical environment. The technical foundation of ARKit lies in two key features: the ability to position the camera in three-dimensional space and the detection of horizontal and vertical planes on non-flat surfaces. To achieve this, ARKit employed a technique known as inertial visual odometry (VIO). This technique was essentially a way to track the device's location in 3D space by combining motion sensor data with information from camera-captured frames.

2. Materials and Methods

2.1. Study Design

The pre-marketing clinical research design was prospective, single-center, and non-randomized, with a single arm of subjects. The objective of the clinical investigation was to collect skin lesion area data by the clinicians to validate the performance and safety of the developed research software by comparing the results obtained between four measurement methods on a total of 65 subjects. The four measurement methods used were the traditional Kundin method of calculation [19,20], digital planimetry with Adobe Photoshop [21], and the developed research software with two different modalities.

2.1.1. Traditional Method

The Kundin method involved using a ruler or a similar measuring device to determine the length and the width in centimeters of a wound at its longest and widest points, respectively (Figure 1). The wound's surface area was calculated by clinicians in square centimeters and adjusted using Kundin's constant (0.785) (Equation (1)).

$$Area = Width \times Height \times 0.785, \tag{1}$$

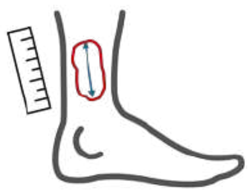

Figure 1. Area measurement with a conventional ruler and Kundin's formula.

2.1.2. Digital Planimetry

Digital planimetry [22] involved capturing an image of the skin lesion and transferring it to a computer. The outline of the skin lesion was traced on the screen by the clinician, using a pointing device and referencing the image to accurately outline the contour. Then, a computer program called Adobe Photoshop [21] was used to compute the number of pixels within the skin lesion outline (Figure 2). To perform this technique, it was necessary to (a) include in the captured image an external calibrator of known measurements, which was located in the same plane as the skin lesion, to serve as a reference; (b) draw the outline of the skin lesion and the external calibrator; (c) calculate the number of pixels within the skin lesion; (d) calculate the number of pixels within the external calibrator; and (e) calculate the area of the skin lesion using the known area of the external calibrator as a reference.

This is an accurate and highly reliable method that does not require contact with the skin lesion, and in this research, it was used as a gold standard for verifying the results obtained with the research software. However, digital planimetry is not used in conventional practice because it is more time-consuming than ruler measurement and other traditional methods [22,23].

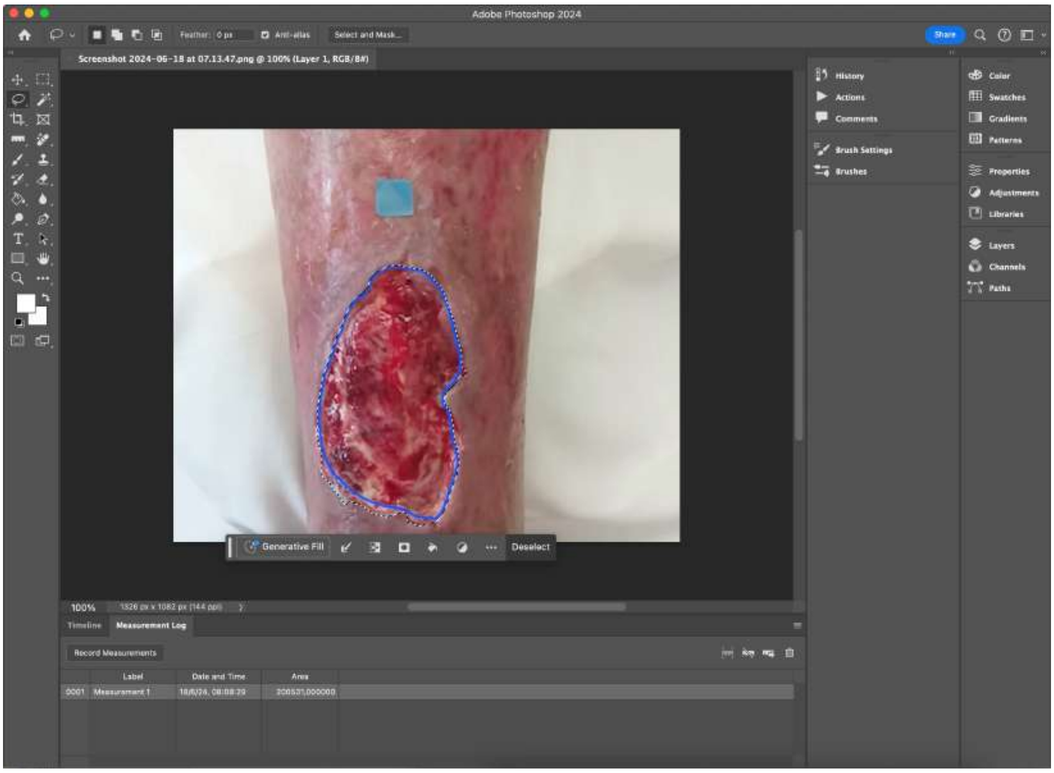

Figure 2. Measuring the area of a skin lesion by digital planimetry with Adobe Photoshop.

2.1.3. Measurement Algorithm Solution

The proposed method for automatically calculating the wound area involved using a software application that was utilized in combination with mobile computing platforms (iPad and iPhone). Software algorithms allowed for the automatic measurement of the wound area when a picture was taken with the mobile application [24].

The software supported two measurement methods. The first method (Figure 3) required the use of an external calibrator consisting of a biocompatible plastic adhesive with known dimensions. The methodology consisted of several steps:

- Image Preprocessing: Initially, the image was converted from RGB to HSV color space. A filter was applied to isolate a range of colors containing various shades of blue, producing a mask with only the pixels within the specified range.
- Contour Detection: All object contours within the mask were identified. Only those with a square shape—meaning twice as long as wide, like our calibrator—were selected.
- Shape Approximation: An artificial vision function was used to approximate all contours to square shapes. For the marker, which had a fairly regular shape, the function searched for the regular shapes that most closely matched the real contour by comparing the areas.
- Shape Selection: If multiple square shapes were detected, the function retained the one with the shape closest to a rectangle. This was determined by the smallest difference between the object's area after applying the function and the actual area initially detected with the mask.
- Calibrator Segmentation and Pixel Calculation: Using the previously detected and approximated shapes, the calibrator was segmented. The number of pixels comprising the calibrator was then calculated. This calibrator served as a scale in the image to compare its proportions with those of the wound.
- Wound Delineation: The outline of the skin lesion was traced on the screen by the clinician, using a pointing device and referencing the image to accurately outline the contour. Once the wound was manually delineated by the specialist, the number of pixels comprising the wound bed was calculated.

Figure 3. Placement of the adhesive calibrator (2D) and image capture for automatic calculation of the area by the software.

The second method (Figure 4) did not require the use of any physical calibrator; instead, it utilized AR as the calibration method through a specific application developed with Apple technology and leveraging ARKit functionality.

In our research, we utilized this functionality to accurately project two virtual points onto the scene captured by the camera. This projection allowed us to precisely calculate the real distance in centimeters between these virtual points, providing crucial reference markers for subsequent area calculation.

Subsequently, the application allowed the delineation of the wound contour to be calculated, taking into account the reference points, which provided the actual calculation in square centimeters (cm^2). The reference points were identified in the image using computer vision techniques similar to those used for detecting the adhesive calibrator. In this case, the points were white and round, and since there were two dots of the same size in each image, we searched for pairs of dots with identical sizes. Once identified, the shortest distance between them was measured to establish a reference in equivalent pixels.

Figure 4. Selection of two points using AR capabilities (3D) and image capture.

In both methods, wound segmentation was performed manually by having the clinician draw the contour. To ensure measurement repeatability, it was necessary to account for the variability in the clinician's manual delineation across different images.

2.2. Patients

The target population consisted of adult patients (>=18 years old) with one or more skin wounds meeting certain criteria: the wound had to be measurable with a 15-centimeter ruler and had to be located in an area that was visible and not obscured. Additionally, it could not have been a carcinoma or other skin lesion presenting malignancy, and it must not have had excessive exudate that could obscure part of the lesion and its contour. The patient was required to hold still for at least 10 seconds in order to take a picture of the lesion.

Patients were excluded if they were currently participating in another clinical investigation, pregnant, or breastfeeding. Additionally, those with medical, social, or psychological conditions that could have limited the subject's ability to participate in the research study were also excluded.

The patients signed an informed consent form. Subsequently, patient screening was performed to determine whether the subject was included or not included in the study. Pictures were taken and no research follow-up was required for any of the patients in the study (Figure 5).

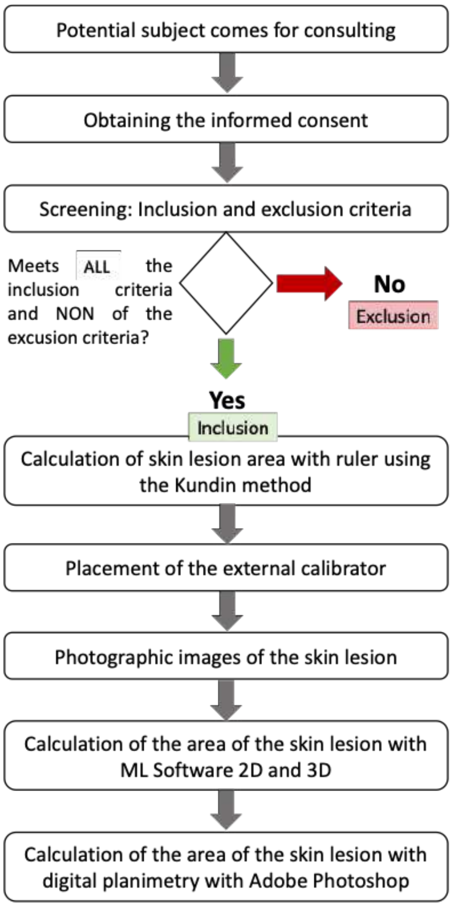

Figure 5. Protocol of the study.

2.3. Sample Collection

The study required a single visit to the center by the subject. The visit consisted of preparing the subject by placing him/her in a suitable and comfortable position, covering the sensitive areas with a cloth, and cleaning the skin lesion with physiological saline solution. The room was conditioned with adequate lighting. Since the position of the light was a key factor in the quality of the image, in order to eliminate shadows in the area to be photographed, the image was not taken under a direct light source. The flash was avoided, and a frontal focus of the light was achieved.

Once the patient had given informed consent, a trained nurse measured the wound area of patients using an application installed on two different iPhone models and an iPad, resulting in three images: one taken with an iPhone (photo A), another with a different iPhone (photo B), and the last with an iPad (photo C). This approach was used to assess not only the agreement of the area measurement results with the software but also the reproducibility and repeatability.

The procedure was as follows (Figure 6): first, the wound's area was measured by the researcher using a ruler and the Kundin method. Following, the researcher placed an external calibrator in the same plane as the skin lesion on intact skin at a minimum distance of 2 cm from the patient's skin lesion and, prior to taking the photograph, the researcher selected two reference points on the screen of the mobile device so that both the 2D and 3D versions of the software would have a reference for calculation. Then, three images of the same wound were taken with the devices using the application, which automatically calculated the area. The calculation of the area with the digital planimetry was performed once the patient left the consultation room. The principal investigator loaded image A of the patient's wound into a computer and, using Adobe Photoshop, manually selected both the adhesive caliper and the wound. The program calculated the pixels of each of the objects, and the principal investigator noted the results to make the subsequent calculation of the wound area.

Different nurses, trained in the use of the software, participated in the study. They independently took the photos and performed the measurements on their respective patients under the supervision of the principal investigator. Subsequently, the principal investigator calculated the gold standard measurements for all samples using digital planimetry.

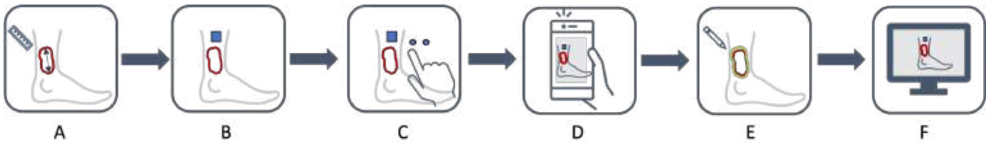

Figure 6. From left to right: (**A**) area measurement with Kundin method, (**B**) placement of adhesive calibrator, (**C**) selection of two virtual points, (**D**) image capture with two iPhones and an iPad, (**E**) delineation of the wound contour, (**F**) area measurement with digital planimetry.

2.4. Sample Size

The main parameter was the calculation of the skin lesion area, measured in square centimeters. According to the study protocol, up to a maximum of 65 patients could have been recruited. Additionally, the study allowed up to 3 skin lesions located in the same patient, so the number of patients recruited was lower because in certain subjects, more than one skin lesion was included. This sample size was calculated based on the primary objective of the study, which was to compare the concordance of skin lesion area calculation in patients in a standard hospital care setting, with two standard practice methods (ruler measurement and digital planimetry), and calculation with the research software.

According to statistical rationale, a sample size of 55 subjects or evaluable skin lesions was defined because a subject could have had more than one evaluable skin lesion; the maximum is to measure 3 skin lesions per subject, with 3 methods of calculating area per skin lesion with a power (1-β) of 90% to detect an intraclass correlation coefficient of 0.90 under the alternative hypothesis with a 95% confidence Interval when the intraclass correlation coefficient under the null hypothesis is 0.8 using an F-test with an α significance level of 0.05 [25]. Anticipating a 15% rate of non-assessable measurements (e.g., due to patients with incomplete records), a sample size of 65 skin lesions was needed.

2.5. Statistics

All relevant statistical analyses were conducted on the variables collected in the study's e-CRF (Electronic Case Report Form) to fulfill the study objectives. Descriptive statistics were utilized to summarize all the data collected, including mean, standard deviation, median, quartiles, minimum, and maximum for continuous variables, as well as absolute and relative frequencies for categorical variables. Valid and missing cases were accounted for in each analysis.

Confidence intervals for the means were computed using the standard normal distribution, while the binomial distribution was employed for proportions.

Data analysis was performed using the SAS statistical package (version 9.4 or higher), with confidence intervals of the intraclass correlation coefficients calculated using R. The statistical method employed for the primary objectives was the Intraclass Correlation Coefficient (ICC), a widely accepted statistical measure for concordance studies involving continuous variables [26].

The ICC has been extensively utilized in the literature for evaluating reliability, intra-operator precision (repeatability), and inter-operator precision (reproducibility) of medical devices [26–28]. Notably, it has been frequently employed in accuracy studies, particularly in assessing the performance of skin lesion area calculation applications similar to the one utilized in this study [22,29,30].

Thus, the ICC was utilized to evaluate the agreement between different methods of skin lesion area measurement, as well as the repeatability and reproducibility of the application across different mobile devices in our research.

3. Results

In accordance with the statistical analysis plan, the unit of analysis was the skin lesion, except for the analysis of usability and also for protocol deviations and reported device deficiencies, where the unit of analysis was the investigator and the patient, respectively. Three populations of analysis were considered:

- Eligible population: Lesions belonging to patients who have signed the informed consent and meet the selection criteria.
- Evaluable population: Eligible lesions that have been measured by the three methods of the study (Kundin, software research measurement, and digital planimetry).
- Population by protocol: Evaluable lesions belonging to patients without major protocol deviations and who have not had deficiencies in the measurement devices. A product deficiency is considered as any inadequacy in the identity, quality, durability, reliability, safety, or efficacy of an investigational product, including its malfunction, errors in use, or the inadequacy of the information provided by the manufacturer according to MDR Regulation 2017/745 [31].

The number of skin lesions measured during the study was 67 (Figure 7). Subsequently, two skin lesions were withdrawn because they were neoplastic skin lesions and the fourth lesion measured from a patient. Therefore, the total number of skin lesions measured, excluding the subjects withdrawn from the study, was 65. The number of patients in whom more than one skin lesion was measured was 17. Thus, the number of patients meeting the eligibility criteria recruited was 41.

There was no follow-up period for patients within the clinical research. Considering that the software application is not an implantable medical device, the risks were deemed acceptable, and the investigator did not use the results of the software for the diagnosis or management of the patients. The usual clinical practice treatment was sufficient for the care of the subjects after the clinical investigation, and no specific medical care was required after the skin lesion area measurement activities performed on the subjects. Subjects participating in the study continued to be treated according to their pathology, in accordance with standard clinical practice.

During the study monitoring visits, 21 minor deviations from the protocol were detected. The corresponding forms were filled out with the reason for the deviation along with corrective actions. No major deviations were reported throughout the study.

Seven device deficiencies were reported under investigation. All of them were related to errors in the use of the device. The research team did not correctly follow the Instructions for Use of the device in relation to the conditioning of the brightness in the room to avoid the appearance of glare, shadows, or flashes in the photograph from which the software must calculate the area of the skin lesion. In these cases, it was observed that the luminosity of the photograph was low. As a consequence, the values obtained for the areas calculated from the poorly illuminated photographs did not correspond to the real values and were often 0.

The analysis of the evaluable population and the correct protocol population, which excludes skin lesions associated with major protocol deviations and device deficiencies, was performed. Considering the exclusion of the 7 lesions associated with device deficiencies, the study includes a total of 58 skin lesions corresponding to 38 patients. Considering that no major protocol deviations were reported, no skin lesions were excluded for this reason.

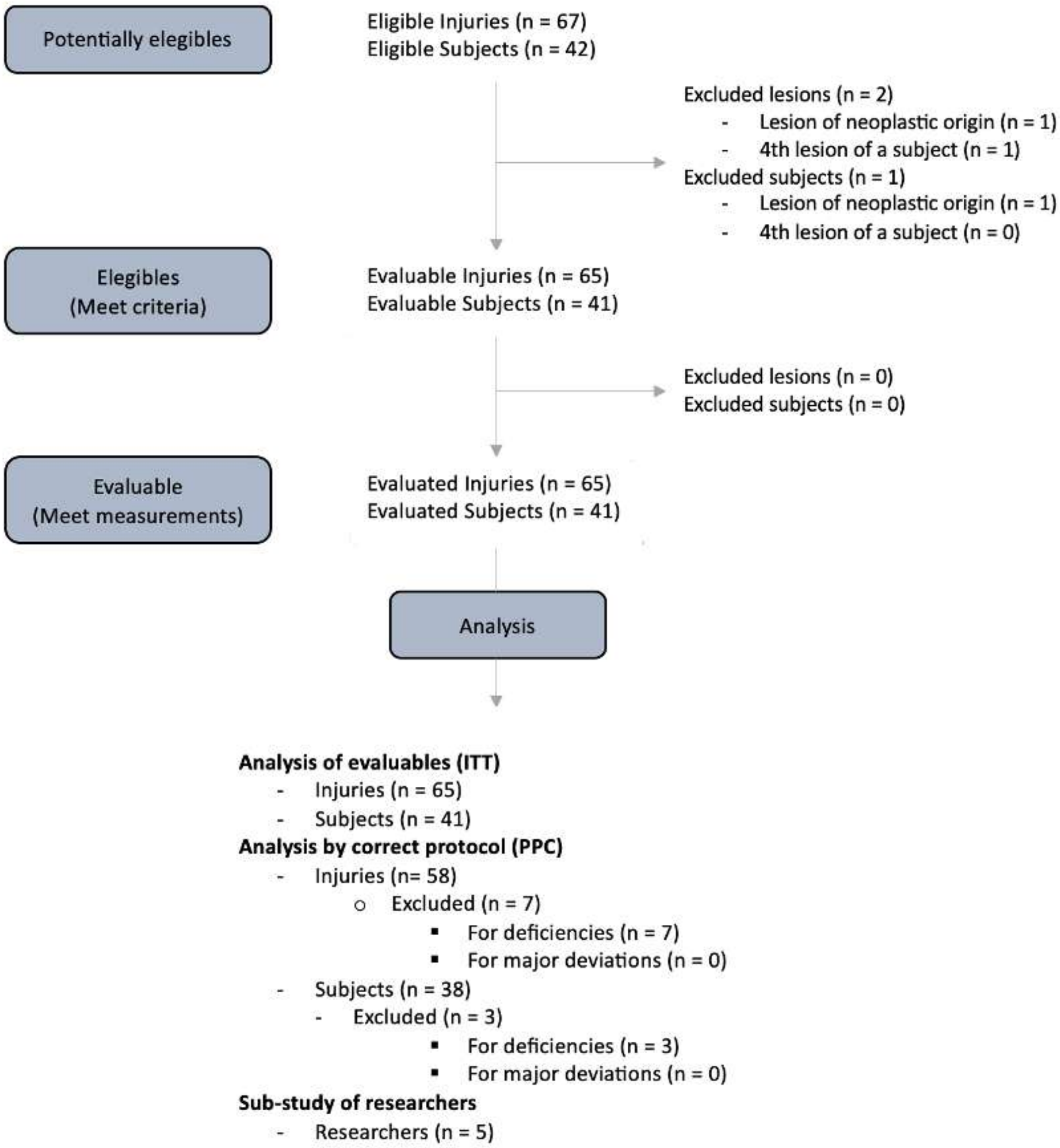

Figure 7. Flow of the resulting sample.

3.1. Demographic Results

The demographic data of the 41 patients included in this clinical study are summarized in Table 1. The mean age of the patients was 78.39 years, ranging from 46 to 96 years. The proportion of females and males was similar, 46.34% and 53.66%, respectively, totaling 19 females and 22 males. The average weight of the patients was 70.93 kg, ranging between 48 and 91 kg.

Regarding mobility, 41.46% of the patients had good mobility, while approximately 29.27% had reduced mobility, and another 29.27% had no mobility or were bedridden. Most patients did not smoke or consume alcohol; only 14.63% (6/41 patients) smoked and 17.07% (7/41 patients) consumed alcohol.

In terms of skin type, 92.68% of the patients had class III skin type, characterized as darker white. 4.88% had skin type II, described as white skin with red or blond hair and blue, hazel, or green eyes. Finally, 2.44% had skin type VI, characterized by dark brown or black skin.

Table 1. Descriptive table of the demographic data of the evaluable population.

Parameter	Feature	Patient
Age (years)	Average Standard Deviation (SD)	78.39 (±11.65)
	Median [P25, P75]	81.00 [72.00, 85.00]
	[Min, Max]	[46.00, 96.00]
Sex	Man	22 (53.66%)
	Woman	19 (46.34%)
Mobility Status	Good mobility	17 (41.46%)
	Reduced mobility	12 (29.27%)
	No mobility/bedridden	12 (29.27%)
Smoker	Yes	6 (14.63%)
	No	35 (85.37%)
Alcohol consumption habit	Yes	7 (17.07%)
	No	34 (82.93%)
Skin type (Fitzpatrick Classification)	I: Pale white skin; red or blond hair; blue/green eyes; freckling	0
	II: Pale white skin; red or blond hair; blue, hazel or green eyes	2 (4.88%)
	III: Darker white; any eye or hair color	38 (92.68%)
	IV: Light brown skin color	0
	V: Brown skin color	0
	VI: Dark brown or black skin color	1 (2.44%)

3.2. Measurement Results

Below is a box plot illustrating the distributions of the area values obtained with each method of the study (Figure 8). It can be observed that most of the calculated areas were distributed within similar ranges. The interquartile ranges of the areas measured with the Kundin method (KUNDIN) and with the software methods from photo C (EU3DFC and EU2DFC) and photo B (EU3DFB and EU2DFB) were slightly larger than those obtained with the software methods from photo A (EU3DFA and EU2DFA). Smaller interquartile ranges were associated with area data calculated with digital planimetry (PLADIG).

The means obtained, represented by different color symbols in each box, ranged from 9.72 (EU2DFC) to 11.04 (EU3DFB) among all methods. On the other hand, the medians, corresponding to the horizontal line within each box, ranged between 3.40 (KUNDIN) and 4.76 (EU3DFB). The upper and lower horizontal lines extending the boxes represent the maximum and minimum values that are within limits defined by +1.5 × P75 and −1.5 × P25, respectively. Several methods had minimum values of 0 or close to 0. In the case of software methods, this was due to device deficiencies associated with operating errors resulting from low luminosity at the time of taking the photographs.

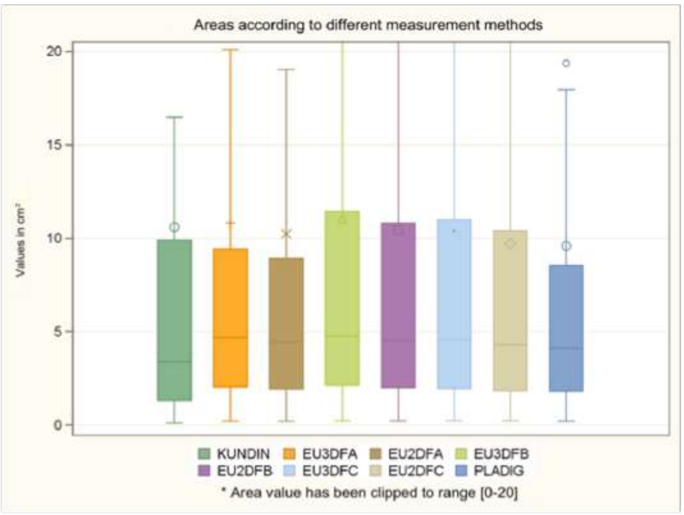

Figure 8. Box plot showing the area values (cm^2) obtained with the different study methods in the evaluable population. For research software, the area values obtained from the different photographs of the skin lesions are also included. The outliers are represented as geometric shapes outside the boxes. The mean is represented as a solid line inside the box.

3.3. Concordance Analysis

The Intraclass Correlation Coefficient (ICC) obtained when assessing the agreement between the methods used to calculate the area of skin lesions in the study is presented in Table 2. According to the protocol, the values of the skin lesion areas were calculated from the same photograph (photo A), corresponding to the first photograph of the lesion taken with the study's iPhone.

The ICC obtained between the software 2D method and the digital planimetry reference method was 0.997, indicating a high degree of agreement. Similarly, the agreement between the software 3D method and digital planimetry was also very high, with an ICC of 0.990. It is noteworthy that the two modes of operation of the software under investigation showed a high degree of agreement, obtaining an ICC of 0.998.

Furthermore, the Kundin method by itself was associated with limitations such as overestimation of the area or lower accuracy for irregular surfaces. Therefore, the agreement between digital planimetry and the Kundin method was also calculated, resulting in an ICC of 0.875. Considering this, it was observed that the agreement between the software 3D and 2D methods and the Kundin method is reflected by ICC values of 0.900 and 0.893, respectively. Once again, it was noted that the agreement between the software 2D and 3D methods and the Kundin method was high.

Table 2. Intraclass correlation coefficients between the different area calculation methods used in the study for the evaluable population. Total of lesions (N): 58.

Parameter	Concordance between Methods ICC *[1] (IC95% *[2])
3D software method iPhone A photo vs. Planimetry	0.990 (0.965, 0.996)
2D software method iPhone A photo vs. Planimetry	0.997 (0.990, 0.999)
3D software method iPhone A photo vs. Kundin Method	0.900 (0.836, 0.939)
2D software method iPhone A photo vs. Kundin Method	0.893 (0.826, 0.935)
3D software method iPhone A photo vs. 2D software method iPhone A photo	0.998 (0.993, 0.999)
Digital Planimetry vs. Kundin Method	0.875 (0.798, 0.924)

*[1] Intraclass Correlation Coefficient: The closer to 1, the greater the degree of agreement. *[2] A 95% confidence interval of the ICC indicates that there is a 95% chance that the true ICC value lies anywhere between the lower and upper limit.

In addition, an analysis of the difference between the measurements was performed using Bland–Altman plots [32]. This graph provides a visual representation of the difference between two measurements on the y-axis and the average of the two measurements on the x-axis. Bland–Altman plots (Figure 9) include the mean difference (red solid line), limits of agreement (red dotted lines) with an approximate 95% CI of the limits of agreement, and the maximum tolerated difference (green dotted lines).

Between photo A measurements from software methods and digital planimetry measurements (Figure 9), the software methods tended to yield slightly higher average measurements of the area of the same lesion (solid red line above the solid blue line), and the larger the lesions to be measured, the slightly higher the software measurements over digital planimetry (linear increasing pattern of the point cloud).

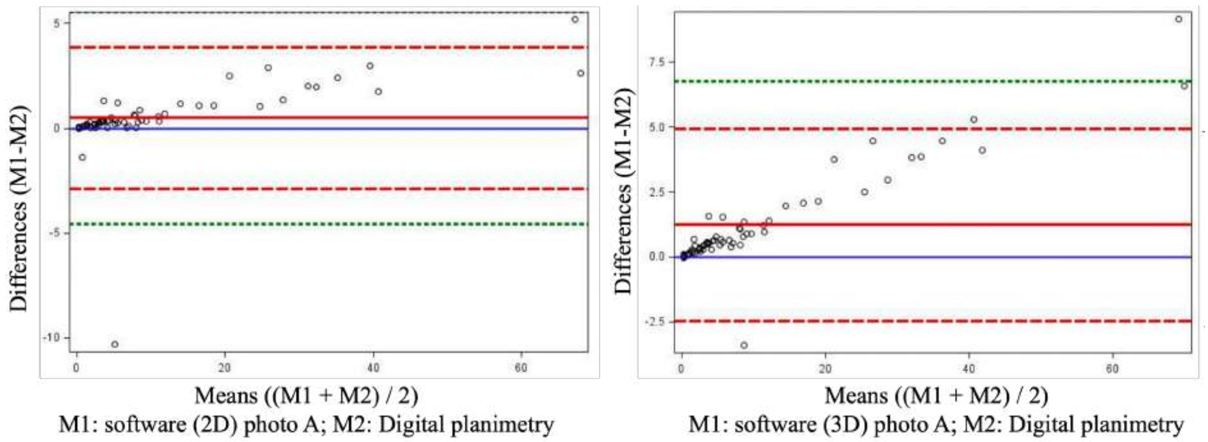

Figure 9. Bland–Altman plots between software 2D method in photo A and digital planimetry (**left**) and between software 3D method in photo A and digital planimetry (**right**). It includes the mean difference (red solid line), limits of agreement (red dotted lines) with an approximate 95% CI of the limits of agreement, and the maximum tolerated difference (green dotted lines).

Between photo A measurements from software methods and Kundin's method (Figure 10), both methods provided similar average area measurements (overlapping solid red and blue lines); although, as the lesion area increased, the differences increased randomly for both methods (the point cloud became similarly distributed above and below zero).

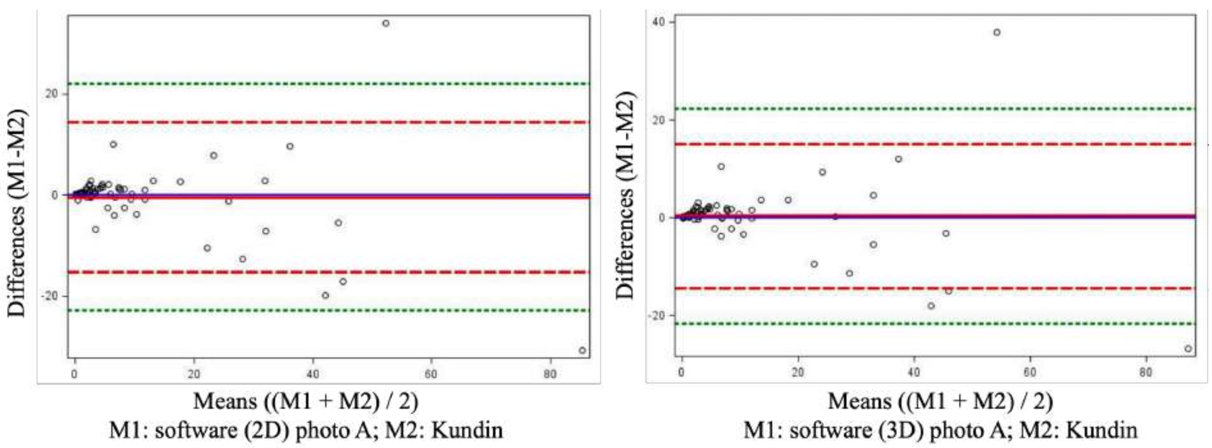

Figure 10. Bland–Altman plots between software 2D method in photo A and Kundin's method (**left**), and between software 3D method in photo A and Kundin's method (**right**). It includes the mean difference (red solid line), limits of agreement (red dotted lines) with an approximate 95% CI of the limits of agreement, and the maximum tolerated difference (green dotted lines).

Between digital planimetry and Kundin, both methods provided similar average area measurements (overlapping solid red and blue lines); although, as the lesion area increased, the differences increased randomly for both methods (the point cloud became similarly distributed above and below zero). Finally, between 3D software photo A measurements and 2D software photo A measurements, both methods offered similar average area measurements (very close solid red and blue lines), with larger measurements in larger lesions (increasing and linear pattern) (Figure 11). In this case, the linear slope indicated the presence of a proportional bias, suggesting that the methods did not agree consistently across the range of measurements in most cases. The two outliers would disprove the hypothesis of a proportional constant error, which could be caused by issues such as a calibration error in a method or a problem with certain constants in an equation when calculating the final results. This increasing linear trend was also slightly noticeable in the plot comparing 3D measurement and digital planimetry (Figure 9).

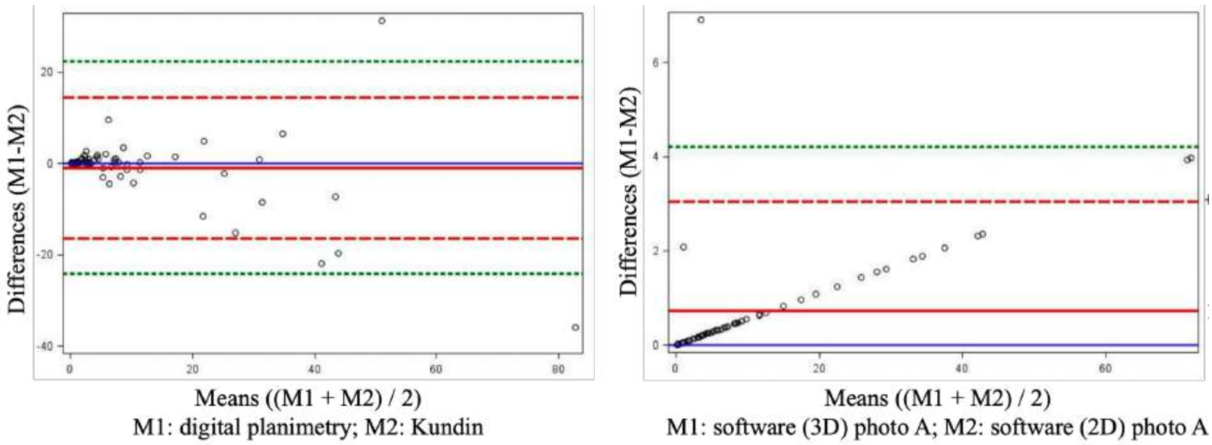

Figure 11. Bland–Altman plots between digital planimetry and Kundin's method (**left**) and between software 3D method in photo A and software 2D method in photo A (**right**). It includes the mean difference (red solid line), limits of agreement (red dotted lines) with an approximate 95% CI of the limits of agreement, and the maximum tolerated difference (green dotted lines).

3.4. Repeatability Analysis

The intra-operator accuracy and repeatability of the area values calculated with the research software from consecutive photographs A and B, taken with the study iPhone, were evaluated. Table 3 presents the obtained results, showing an ICC of 0.996 for the areas calculated with the 3D method of the research software. This indicated a high level of intra-operator repeatability for the 3D method. Similarly, the area values calculated with the 2D method exhibit an ICC of 0.996, signifying an even higher degree of intra-operator repeatability. These findings suggested that both the 3D and 2D methods consistently provide accurate skin lesion area values when calculated from photographs captured by the same mobile device and user at consecutive time points. Thus, the repeatability of both the 3D and 2D research software methods has been demonstrated.

Table 3. Intraclass correlation coefficients obtained to evaluate the repeatability of research software in the evaluable population. Total of lesions (N): 58.

Parameter	Repeatability with Research Software ICC (IC95%)
3D software method iPhone photograph A vs. 3D software method iPhone photo B	0.996 (0.993, 0.997)
2D software method iPhone photograph A vs. 2D software method iPhone photo B	0.996 (0.993, 0.997)

3.5. Reproducibility Analysis

To investigate the inter-device accuracy and reproducibility of the area values calculated with the research software, photographs A taken with the study iPhone and photograph C taken with the iPad were evaluated. Table 4 illustrates the obtained results, indicating an ICC of 0.980 between the areas calculated from photographs A (iPhone) and C (iPad) using both the 3D and 2D software methods. These findings demonstrated a very high degree of inter-device reproducibility for both modes of operation of the software. Thus, it is observed that the 3D and 2D methods consistently provide accurate skin lesion area values regardless of the mobile device used for photography, whether they are derived from photographs taken by an iPhone or an iPad. Consequently, the inter-device reproducibility of the research software's 3D and 2D methods has been established.

Table 4. Intraclass correlation coefficients obtained to evaluate the reproducibility of research software using two different mobile devices: iPhone and iPad in the evaluable population. Total of lesions (N): 58.

Parameter	Reproducibility between Devices ICC (IC95%)
3D software method iPhone photograph A vs. 3D software method iPad photo C	0.980 (0.967, 0.988)
2D software method iPhone photograph A vs. 2D software method iPad photo C	0.980 (0.967, 0.988)

4. Discussion

As observed in the literature, measuring the area of a wound is a crucial parameter for monitoring wound healing progress [6–10]. Currently, new technologies and mobile applications have simplified wound monitoring through image recording. Traditional measurement methods, such as using a ruler, are functional due to their low cost and availability in any hospital. However, these methods are often inaccurate and time-consuming. An optimized approach leverages technology to automatically calculate wound area during imaging for follow-up.

In this study, we proposed two methods for calculating wound area: a 2D method utilizing an adhesive calibrator, which the clinician places near the wound without contact, followed by image capture, and a 3D method employing AR technology, allowing the clinician to capture an image and select two reference points on the mobile device screen without patient contact. Both methods are integrated into a single application functional on iOS devices.

The clinical study demonstrated the accuracy, repeatability, and reproducibility of the 2D and 3D methods for area measurement, compared to the traditional Kundin and digital planimetry methods.

Box plot analysis revealed consistent distributions of area values across different methods. The correlation indices were higher when comparing the 2D and 3D methods with digital planimetry (ICC = 0.997 for 2D and ICC = 0.990 for 3D) than with the Kundin method (ICC = 0.893 in 2D and ICC = 0.900 in 3D). Consistent with previous research emphasizing the accuracy of digital planimetry in wound measurement [33], our developed 2D and 3D methods show higher reliability in wound area calculation than traditional methods.

Bland–Altman plots confirmed the correlation between both methods and digital planimetry and Kundin, although slight proportional biases were observed in measurements, particularly for larger lesions.

Repeatability analysis demonstrated high correlation for the 2D and 3D methods (ICC = 0.996 for both), indicating consistent performance in consecutive photographs using the same method.

Reproducibility analysis confirmed the high accuracy and inter-device reproducibility of the two software methods, with ICC values indicating strong agreement between measurements made with different devices (ICC = 0.980 for both methods comparing photo A with photo C).

Comparative analysis with other market-available methods indicated a higher correlation index with the gold standard in our study. For instance, NDKare, when compared with Visitrak and WoundVue, achieved a maximum ICC of 0.991 [15], which is lower than the ICC between our 2D method and digital planimetry. However, the ICC values between our 2D method and Kundin were lower, while the 3D method demonstrated higher ICC values compared to Kundin (ICC = 0.998). In another study comparing AreaMe and SilhouetteMobile with Visitrak, box plot analysis showed that the distribution ranges of the AreaMe and SilhouetteMobile methods were more distant from traditional methods compared to our study's software methods versus digital planimetry and Kundin [16].

Therefore, it could be stated that our study provided a solid validation of a novel tool that provides two ways to measure wound area with high accuracy, whereas other studies only present one method with one particular technology. Furthermore, comparisons with the results of other studies underlined the reliability and performance of the methods in our study, which places them as valuable tools for wound assessment and monitoring in clinical practice.

However, as a point to improve, the methods developed in our study are functional and tested only on iOS devices. In future developments, it is intended to deploy and test it on Android devices.

5. Conclusions

The conducted pre-market clinical investigation validated the performance, safety, and usability of the ML method as an effective tool for measuring skin lesions. These findings underscore the potential of the software to enhance clinical practice by providing accurate and efficient measurement of skin lesions.

Overall, the software-based methods for calculating skin lesion area proved to be accurate, reliable, and safe for clinical use, suggesting their potential to enhance patient care and clinical decision-making in skin lesion management.

Author Contributions: Conceptualization, L.C.-L. and D.R.-J.; methodology, L.C.-L. and D.R.-J.; software, L.C.-L. and D.R.-J.; validation, L.C.-L. and D.R.-J.; investigation, M.d.M.M.-E.; writing—original draft, L.C.-L. and D.R.-J.; writing—review and editing, M.d.M.M.-E., R.R.-B., and S.G.-C.; supervision, M.d.M.M.-E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Spanish Agency of Medicines and Health Products (AEMPS) (protocol code CIP-EUCLIDES-01; date of approval 23 November 2022). The favorable opinion of the CEIm for the conduct of the study, with Protocol internal code 2022/79-ENF-HUSC Version 2.0 20 September 2022 and Patient Information Sheet/Informed Consent Version 2.0 20 September 2022, by the Principal Investigator of the Department of Nursing at the Hospital Universitari Sagrat Cor was obtained on 3 October 2022. Clinical trial information is registered and available at [34].

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are unavailable due to privacy or ethical restrictions.

Acknowledgments: We gratefully acknowledge the participation of the healthcare professionals from Hospital Universitari Sagrat Cor, who captured and recorded the patient data, making the development of this clinical trial possible. This paper was presented as a poster at the 6th International Conference on Advances in Signal Processing and Artificial Intelligence (ASPAI' 2024), held on 17–19 April 2024, in Funchal, Madeira Island, Portugal [35].

Conflicts of Interest: The authors declare the following potential conflicts of interest with respect to the research of this article: L. Casanova and D. Reifs are employees of Skilled Skin S.L. (Clinicgram's manufacturer), the organization that provided the funding (Industrial Doctorate Program) and

resources for this research. While the company supported the research project, it had no direct involvement in the study design, data collection, analysis, interpretation of data, or the decision to publish the results. To mitigate any potential bias, an independent Contract Research Organization (CRO), MED IVD HEALTHTECH S.L., was contracted to oversee the study design, data collection, and initial data analysis. The project was conducted at Hospital Universitari Sagrat Cor-Quironsalud, ensuring a neutral and reputable environment for the research. The study protocol was validated by the ethics committee at Quironsalud, and the study received approval from the Spanish Agency of Medicines and Medical Devices (AEMPs), ensuring compliance with all regulatory and ethical standards. The authors affirm that this affiliation has not influenced the objective and impartial nature of the research presented in this paper and commit to upholding the highest standards of transparency and integrity in their work.

References

1. Marijanović, D.; Filko, D. A systematic overview of recent methods for non-contact chronic wound analysis. *Appl. Sci.* **2020**, *10*, 7613 [CrossRef]
2. González de la Torre, H.; Soldevilla-Ágreda, J.; Verdú-Soriano, J. Clinical Practice: The status of wound care units in Spain. *Wounds Int.* **2017**, *8*, 28–33.
3. Lindholm, C.; Searle, R. Wound management for the 21st century: Combining effectiveness and efficiency. *Int. Wound J.* **2016**, *13*, 5–15. [CrossRef] [PubMed]
4. Wilkinson, H.N.; Hardman, M.J. Wound healing: Cellular mechanisms and pathological outcomes. *Open Biol.* **2020**, *10*, 200223. [CrossRef] [PubMed]
5. Sen, C.K. Human Wound and Its Burden: Updated 2020 Compendium of Estimates. *Adv. Wound Care* **2021**, *10*, 281–292. [CrossRef]
6. Blasco Vera, M.Á.; Aunés García, L.; Blanes Ortí, P.; Ramos Romero, I.; Hernández Sanfelix, A. Sistemas de medición de heridas. *Rev. Enferm. Vasc.* **2019**, *2*, 17–21. [CrossRef]
7. Gethin, G. The importance of continuous wound measuring. *Wounds* **2006**, *2*, 60–68.
8. Nicholas, E. Wound Assessment Part -I: How to measure a wound. *Wound Essentials* **2015**, *10*, 51–55.
9. Mohíno, A.G.P.G.; Moratalla, M.S.; López, C.C.; García, L.R.; López, N.P.; Sáez, L.H.; Luengo, L.G.; Carrilero, R.N.; Álvarez, T.G.; Fernández, F.R.; et al. Guia de Prevención y Manejo de Úlceras por Presión y Heridas Crónicas, Grupo de Trabajo de Ulceras por Presión y Heridas cróNicas. Gerencia de atención integrada de Albacete, SESCAM. 2015. Available online: https://www.chospab.es/enfermeria/ulceras_heridas/prevencion_manejo.htm (accessed on 18 May 2024).
10. Wynne, M. National Wound Management Guidelines. The Office of Nursing and Midwifery Services Director, Clinical Strategy and Programmes Division, Dr Steevens' Hospital: Dublin 8. 2018. Available online: www.hse.ie/eng/about/who/onmsd/practicedevelopment/WoundManagement/ (accessed on 18 May 2024).
11. Dowsett, C.; von Hallern, B. The Triangle of Wound Assessment: a holistic framework from wound assessment to management goals and treatments. *Wounds Int.* **2017**, *8*, 34–39.
12. Ruiz, E.A.; López, J.R.; Vicente, J.M.S. Actualización en la validez de las escalas de evaluación de la evolución de heridas. *Heridas Cicatrización* **2021**, *11*, 15–21.
13. Zahia, S.; Garcia-Zapirain, B.; Elmaghraby, A. Integrating 3D model representation for an accurate non-invasive assessment of pressure injuries with deep learning. *Sensors* **2020**, *20*, 2933. [CrossRef] [PubMed]
14. Pena, G.; Kuang, B.; Szpak, Z.; Cowled, P.; Dawson, J.; Fitridge, R. Evaluation of a Novel Three-Dimensional Wound Measurement Device for Assessment of Diabetic Foot Ulcers. *Adv. Wound Care* **2020**, *9*, 623–631. [CrossRef]
15. Kuang, B.; Pena, G.; Szpak, Z.; Edwards, S.; Battersby, R.; Cowled, P.; Dawson, J.; Fitridge, R. Assessment of a smartphone-based application for diabetic foot ulcer measurement. *Wound Repair Regen.* **2021**, *29*, 460–465. [CrossRef] [PubMed]
16. Foltynski, P.; Ladyzynski, P.; Wojcicki, J.M. A new smartphone-based method for wound area measurement. *Artif. Organs* **2014**, *38*, 346–352. [CrossRef]
17. Mohammed, H.T.; Bartlett, R.L.; Babb, D.; Fraser, R.D.; Mannion, D. A time motion study of manual versus artificial intelligence methods for wound assessment. *PLoS ONE* **2022**, *17*, 1–14. [CrossRef]
18. Apple Inc. Understanding World Tracking—Discover Features and Best Practices for Building Rear-Camera AR Experiences. Available online: https://developer.apple.com/documentation/arkit/arkit_in_ios/configuration_objects/understanding_world_tracking (accessed on 2 May 2024).
19. Kundin, J.I. A new way to size up a wound. *Am. J. Nurs.* **1989**, *89*, 206–207. [CrossRef]
20. Restrepo-Medrano, J.C.; Verdú, J. Measure healing in pressure ulcers. What do we have? *Gerokomos* **2011**, *22*, 35–42.
21. Adobe Inc. Adobe Photoshop. Available online: https://www.adobe.com/products/photoshop.html (accessed on 6 March 2019).
22. Jørgensen, L.B.; Sørensen, J.A.; Jemec, G.B.; Yderstræde, K.B. Methods to assess area and volume of wounds—A systematic review. *Int. Wound J.* **2016**, *13*, 540–553. [CrossRef] [PubMed]
23. Hasatsri, S.; Aramwit, P. Nontraditional Methods to Evaluate Wound Healing. *Dermatol. Surg.* **2017**, *43*, 342–350. [CrossRef]
24. Reifs, D.; Casanova-Lozano, L.; Reig-Bolaño, R.; Grau-Carrion, S. Clinical validation of computer vision and artificial intelligence algorithms for wound measurement and tissue classification in wound care. *Inform. Med. Unlocked* **2023**, *37*, 101185. [CrossRef]

25. Walter, S.D.; Eliasziw, M.; Donner, A. Sample size and optimal designs for reliability studies. *Stat. Med.* **1998**, *17*, 101–110. [CrossRef]
26. Liljequist, D.; Elfving, B.; Roaldsen, K.S. Intraclass correlation—A discussion and demonstration of basic features. *PLoS ONE* **2019**, *14*, e0219854. [CrossRef]
27. Lin, L. Overview of agreement statistics for medical devices. *J. Biopharm. Stat.* **2008**, *18*, 126–144. [CrossRef] [PubMed]
28. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [CrossRef] [PubMed]
29. Sigam, P.; Denz, M. Reliability and accuracy of wound surface measurement using mobile technology. *J. Telemed. Ehealth* **2015**, *1*, 38–41.
30. Khac, A.D.; Jourdan, C.; Fazilleau, S.; Palayer, C.; Laffont, I.; Dupeyron, A.; Verdun, S.; Gelis, A. mHealth app for pressure ulcer wound assessment in patients with spinal cord injury: Clinical validation study. *JMIR mHealth uHealth* **2021**, *9*, e26443. [CrossRef] [PubMed]
31. The European Parliament and the Council of the European Union. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EE. *Off. J. Eur. Union* **2017**.
32. Doğan, N.Ö. Bland-Altman analysis: A paradigm to understand correlation and agreement. *Turk. J. Emerg. Med.* **2018**, *18*, 139–141. [CrossRef]
33. Ibraheem, W.I.; Bhati, A.K.; Hakami, N.A.; Alshehri, A.D.; Wadani, M.H.M.; Ageeli, F.M.E. Comparison of Digital Planimetry and Ruler Methods for the Measurement of Extraction Socket Wounds. *Medicina* **2023**, *59*, 135. [CrossRef]
34. ClinicalTrials.gov. Clinical Investigation EUCLIDES-01 for the Calculation of the Area of Skin Lesions. Available online: https://www.clinicaltrials.gov/study/NCT05729334?term=EUCLIDES&rank=1 (accessed on 10 May 2019).
35. Casanova, L.; Reifs, D.; Reig, R.; Grau, S. Assessing Chronic Wound Area Measurement with Machine Learning Techniques in a Single Center, Non-randomized Controlled Clinical Trial. In *Advances in Signal Processing and Artificial Intelligence: Proceedings of the 6th International Conference on Advances in Signal Processing and Artificial Intelligence (ASPAI' 2024)*; Sergey, Y.Y., Ed.; IFSA: Freiburg im Breisgau, Germany, 2024; pp. 121–123.