Visual assessment of digital ulcers in systemic sclerosis analysed by eye tracking: implications for wound assessment

T. Moser[1], Q. Lohmeyer[2], M. Meboldt[2], O. Distler[1], M.O. Becker[1]

[1]Department of Rheumatology, University Hospital Zurich;
[2]Product Development Group Zurich, Department of Mechanical and Process Engineering, ETH Zurich, Switzerland.

Thomas Moser, MD
Quentin Lohmeyer, PhD
Mirko Meboldt, PhD, Prof.
Oliver Distler, MD, Prof.
Mike Oliver Becker, MD, MSc

Please address correspondence to:
Mike Oliver Becker,
Department of Rheumatology,
University Hospital Zurich,
Gloriastrasse 25,
8091 Zurich, Switzerland.
E-mail: mikeoliver.becker@usz.ch

Received on May 7, 2020; accepted in revised form on June 16, 2020.

Clin Exp Rheumatol 2020; 38 (Suppl. 125): S137-S139.

Key words: systemic sclerosis, skin ulcer, eye tracking, gaze pattern analysis

Competing interests: O. Distler has had consultancy relationships and/or received research funding from AbbVie, Actelion, Acceleron Pharma, Amgen, AnaMar, Baecon Discovery, Blade Therapeutics, Bayer, Boehringer Ingelheim, Catenion, Competitive Corpus, Drug Development International Ltd, CSL Behring, ChemomAb, Ergonex, Galapagos NV, Glenmark Pharmaceuticals, GSK, Horizon (Curzion) Pharmaceuticals, Inventiva, Italfarmaco, iQone, iQvia, Kymera Therapeutics, Lilly, Medac, Medscape, Mitsubishi Tanabe Pharma, MSD, Novartis, Pfizer, Roche, Sanofi, Target Bio Science and UCB in the area of potential treatments of scleroderma and its complications. In addition, he has a patent issued mir-29 for the treatment of systemic sclerosis (US8247389, EP2331143).
The other co-authors have declared no competing interests.

ABSTRACT

Objective. *The assessment of digital ulcers (DUs) in systemic sclerosis (SSc) depends crucially on visual aspects. However, little is known about the differences in visual assessment of these wounds between experts and non-experts or medical lay persons (novices). To develop potential training recommendations for the visual assessment of digital ulcers in SSc, we analysed gaze pattern data during assessment of digital ulcers between assessors with different levels of expertise.*

Methods. *Gaze pattern data from 36 participants were collected with a mobile eye tracking device. 20 pictures from digital ulcers of SSc patients were presented to each participant. The analysis comprised the scan path, the dwell times (for areas of interest), fixation counts and the entry time for each picture and subject.*

Results. *Most significant differences were found between novices and medically educated groups. Dwell times in the wound area for novices differed statistically significantly from all medically educated groups (1.76s–3.1s less). Above all, novices had lower dwell times in wound margin and wound surrounding and spent more time in other areas (up to 1.92s longer). Correspondingly, they had less fixation points and longer overall scan paths in wound areas. Similar gaze pattern data were generated for medically educated groups.*

Conclusion. *For the first time, we could analyse the visual assessment of digital ulcers in SSc and detected differences according to levels of medical education and expertise. Adequate training on proper interpretation of changes in all wound areas are warranted to improve wound assessment in digital ulcers.*

Introduction

Digital ulcers (DU) are chronic wounds that are characteristic vascular features of systemic sclerosis (SSc) included in the classification criteria (1). Although treatment of DUs relies on their correct assessment, this is hampered by the lack of agreement over: i) definitions of DUs and ii) their grading (by severity of the lesions), although a preliminary consensus has been proposed (2-5).

Our aim was to analyse differences of visual assessment between raters of different experience levels, which has not been done before. Since this is not feasible to perform with patients and actual DUs, we used the eye tracking technology and gaze pattern analysis, which has been successfully used in other disciplines and clinical applications (6).

Patients and methods

We analysed the gaze pattern data from 36 subjects: 9 expert medical professionals (EMP), 8 non-expert medical professionals (NEMP), 9 medical students (medical graduates, MG) and 10 lay persons (novices), definitions are provided in the Supplementary file (and Supplementary Tables S1 and S2). Twenty pictures from digital ulcers of SSc patients were presented to each participant 30 seconds each in a standardised fashion (for details see Suppl. Tables S3-5). This research was undertaken with approval of the local ethics committee: patients were included upon informed consent and participants also consented. The analysis of gaze pattern data was done using a standardised technique with a mobile eye tracking device (see Suppl. file). Analysed parameters encompassed the scan path, dwell times (for areas of interest, AOI), fixation counts and the entry time for each picture and subject. Areas of interest were defined as the wound ground, wound margin and wound surrounding as well as distracting and non-relevant parts of the pictures (agreed upon by two experts, see Suppl. file and Appendix). Statistical analysis/graphing was made using IBM SPSS v. 21 and Graphpad Prism 8.0.

Results

General eye tracking data

Of the 36 participants, 20 gaze pattern datasets were obtained (1 for each picture displayed), 710 gaze pattern datasets could be analysed (Suppl. Table S7). The general eye tracking are displayed in supplementary Tables S9 and S10. The eye tracking ratio was >96% overall. The scanpath length (SL), which is the sum of all saccadic eye movements, was significantly different amongst groups, with novices having a significantly longer scanpath length than other groups ($p<0.001$, Suppl. Table S1). Fixation counts, which count the number a person focuses on an area less than 100 pixels for at least 80ms, ranged between 75.61 fixations per 30 seconds for the expert medical professionals and 76.68 fixations for the novice group. The cumulative duration of all fixations, the fixation duration total (FDT), was lower in the non-expert medical professionals than in the expert medical professional group ($p<0.001$) and the medical student group ($p=0.001$). Novices also had a lower fixation duration total than expert medical professionals ($p=0.031$).

Dwell times

The dwell times, *i.e.* the time that a participant gaze focused on a certain area of interest, were significantly different amongst the groups for the whole wound (W) area (Fig. 1A): Novices spent less time than all other groups in this area (mean 18.62s; SD 5.39s; $p<0.006$). Another significant difference was between the non-expert and expert medical professional group ($p=0.041$). In addition, novices spent significantly less time in the wound margin (WM) and wound surrounding (WS) area ($p<0.05$; Fig. 1A). The time that was lost in the wound area by the novice group was spent in the "Other" (O) area and the "White space" (S) area compared to the other groups, with significant post-tests ($p<0.001$; Fig. 1B).

Fixation counts

There was a significant difference between groups for fixation counts in the wound area with the novices having the lowest fixation ($p<0.000$; Fig. 2A). Similar to dwell times, this was due to differences in the wound margin and wound surrounding area (Fig. 2A). Again, the novice group was different from all other groups ($p<0.05$ for both areas). Concerning the wound margin area, the novices were only significantly different to the expert medical professional and medical student groups ($p<0.004$; Fig. 2B), whereas for the wound surrounding the post-hoc tests showed a significant difference to all other groups ($p<0.026$; Fig. 2B). Again, similar to the dwell times, fixation counts of the novice group were higher in the "other" and "white space" areas ($p<0.000$; Fig. 2B).

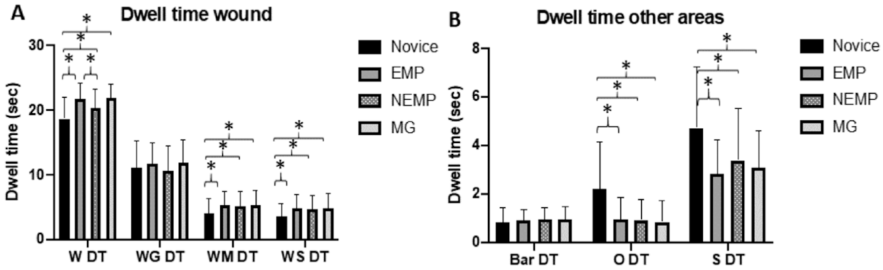

Fig. 1. Dwell times (DT) per subject group (in seconds) for the wound area (W) and sub-areas wound ground (WG), wound margin (WM) and wound surrounding (WS, Fig. 1A). Dwell times for other areas of the picture (Bar, other - O and "white space" - S) are depicted in Fig. 1B. Significance was calculated with Kruskal-Wallis test (W DT, O DT, S DT) or one-way ANOVA (WM DT, WS DT) depending on unequal variances. Significant differences by post-tests ($p<0.05$) are marked with*.

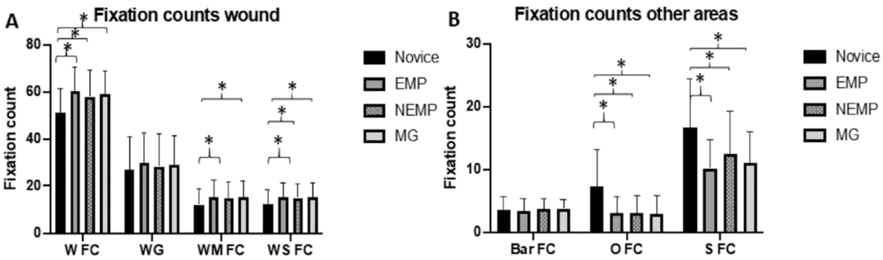

Fig. 2. Fixation counts (FC) per subject group for the wound area (W) and sub-areas wound ground (WG), wound margin (WM) and wound surrounding (WS, Fig. 2A). Fixation counts for other areas of the picture (Bar, other - O and "white space" - S) are depicted in Fig. 2B. Significance was calculated with one-way ANOVA (W FC, WG FC, WM FC) or Kruskal-Wallis (O FC, S FC) depending on unequal variances. Significant differences by post-test ($p<0.05$) are marked with*

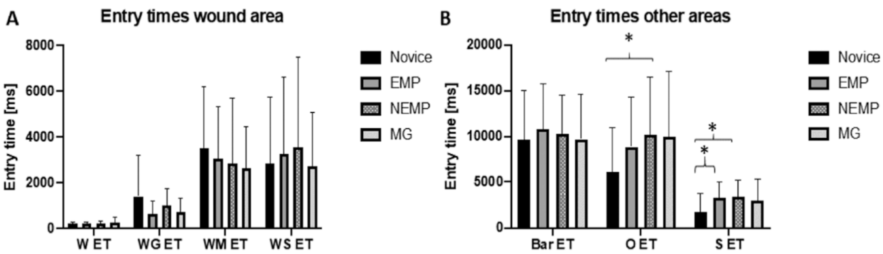

Fig. 3. Entry times (ET) per subject group for the wound area (W) and sub-areas wound ground (WG), wound margin (WM) and wound surrounding (WS, Fig. 3A). Entry times for other areas of the picture (Bar, other - O and "white space" - S) are depicted in Fig. 3B. Significance was calculated by Kruskal-Wallis due to unequal variances. Significant differences by post-test ($p<0.05$) are marked with*.

Entry time

The entry time, *i.e.* the time until a participant gaze first hit a certain area, was lowest for the wound area in all groups (Fig. 3A). The expert medical professionals group hit the wound area first, followed by the medical students and the non-expert medical professionals, there were no other significant differences (Fig. 3A). The expert and non-expert medical professionals as well as the medical students had a tendency to analyse the pictures from centre to periphery. The "other" area was hit earlier by novices than the non-expert medical professionals ($p<0.004$ by post-test) and the "white space" area (mostly unstructured, white areas on the pictures) was also hit earlier by the novices than by the expert or non-expert medical professionals ($p<0.001$ each by post-test; Fig. 3B).

Discussion

For the first time, the eye tracking technology was used to obtain gaze pattern data from assessments of digital ulcers in systemic sclerosis. While novices spend less time and have a lower fixation count observing the wound, especially wound margin and surrounding, they spend more time in other areas of the picture supposedly not relevant for assessment, producing a longer scanpath. The difference was most profound when compared to the expert group, but sometimes also apparent compared to medical students or non-expert medical professionals. This has implications for the training of professionals that treat digital ulcers: a thorough assessment of all wound areas is mandatory, not only focusing on most salient features. In a setting similar to ours, experts and non-experts assessed CT images with cerebrovascular accidents (7): likewise, non-experts paid less attention to the less salient but important features of the CT images. Another study involving trainees and experts in mammography found also that non-experts spent too little time on pathological lesions and too much time observing lesions that were non-pathological, hence having a lower diagnostic accuracy (8).

We observed a trend towards shorter scanpath length and mean fixation duration for medically educated staff, which suggests that taking in relevant information does not necessarily equal a longer time of assessment. This is in accordance with another study, where expert physicians were able to spot relevant imaging patterns earlier and spent more time looking at these patterns (6). With regards to imaging, a systematic review found that the fixation duration increased with the level of expertise (9). Training of inexperienced staff should hence be aimed at a) including all wound areas in the assessment, b) using a systematic approach or an algorithm to avoid stopping the assessment with early pathological findings and c) ignoring distracting information.

Because the analysis of actual patients and ulcers was not feasible in this setting, we used pictures of digital ulcers and this is one limitation of the study. For other chronic wounds, the use of pictures has produced fairly high agreement rates between live assessments and assessments by pictures (10, 11). In the study by Salmhofer *et al.* less experienced junior doctors tended to assess the wound margin and surrounding less than experienced staff. A structured assessment questionnaire, the revised photographic wound assessment tool (revPWAT), found an excellent agreement between bedside and pictorial assessment (12). However, pilot studies investigating the intra- and inter-rater reliability of chronic wound grading by rheumatologist using pictures of DUs found that there is poor inter-rater reliability (2, 3).

Another important limitation of the study is the low number of participants, mainly due to the pilot character of the study. Experts were also involved in defining the areas of interest (AOI). In addition, variability of photographs can be a problem, as demonstrated in a study involving pressure ulcers (13).

In spite of the limitations, we were able to analyse the visual assessment of DUs in SSc patients for the first time. Our results emphasise adequate training in chronic wound care and confirm that training might even enable relative inexperienced staff to support chronic wound care. Proper training could also limit inter-rater variability in the visual assessment of digital ulcers. Standardised assessment of different wound features, with feedback, could eventually improve treatment of DUs. Implementation of assessment algorithms in computer systems or mobile applications might lead to an improved recognition of wound characteristics and be beneficial in routine care or clinical trials.

Acknowledgements

We acknowledge the contribution of the voluntary participants in this study, who received no fees or compensation. This work is part of the SKINTEGRITY project under the umbrella of University Medicine Zurich. Part of these data have been published as a thesis at the Medical Faculty of the University of Zurich, Switzerland.

References

1. VAN DEN HOOGEN F, KHANNA D, FRANSEN J *et al.*: 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. *Ann Rheum Dis* 2013; 72: 1747-55.
2. HUGHES M, ROBERTS C, TRACEY A, DINSDALE G, MURRAY A, HERRICK AL: Does the clinical context improve the reliability of rheumatologists grading digital ulcers in systemic sclerosis? *Arthritis Care Res* (Hoboken) 2016; 68: 1340-5.
3. HUGHES M, TRACEY A, BHUSHAN M *et al.*: Reliability of digital ulcer definitions as proposed by the UK Scleroderma Study Group: A challenge for clinical trial design. *J Scleroderma Relat Disord* 2018; 3: 170-4.
4. BARON M, CHUNG L, GYGER G *et al.*: Consensus opinion of a North American Working Group regarding the classification of digital ulcers in systemic sclerosis. *Clin Rheumatol* 2014; 33: 207-14.
5. SULIMAN YA, BRUNI C, JOHNSON SR *et al.*: Defining skin ulcers in systemic sclerosis: systematic literature review and proposed World Scleroderma Foundation (WSF) definition. *J Scleroderma Relat Disord* 2017; 2: 115-20.
6. BLONDON K, WIPFLI R, LOVIS C: Use of eye-tracking technology in clinical reasoning: a systematic review. *Stud Health Technol Inform* 2015; 210: 90-4.
7. MATSUMOTO H, TERAO Y, YUGETA A *et al.*: Where do neurologists look when viewing brain CT images? An eye-tracking study involving stroke cases. *PLoS One* 2011; 6: e28928.
8. NODINE CF, MELLO-THOMS C, KUNDEL HL, WEINSTEIN SP: Time course of perception and decision making during mammographic interpretation. *AJR Am J Roentgenol* 2002; 179: 917-23.
9. WOOD G, KNAPP KM, ROCK B, COUSENS C, ROOBOTTOM C, WILSON MR: Visual expertise in detecting and diagnosing skeletal fractures. *Skeletal Radiol* 2013; 42: 165-72.
10. HALSTEAD LS, DANG T, ELROD M, CONVIT RJ, ROSEN MJ, WOODS S: Teleassessment compared with live assessment of pressure ulcers in a wound clinic: a pilot study. *Adv Skin Wound Care* 2003; 16: 91-6.
11. SALMHOFER W, HOFMANN-WELLENHOF R, GABLER G *et al.*: Wound teleconsultation in patients with chronic leg ulcers. *Dermatology* 2005; 210: 211-7.
12. THOMPSON N, GORDEY L, BOWLES H, PARSLOW N, HOUGHTON P: Reliability and validity of the revised photographic wound assessment tool on digital images taken of various types of chronic wounds. *Adv Skin Wound Care* 2013; 26: 360-73.
13. JESADA EC, WARREN JI, GOODMAN D *et al.*: Staging and defining characteristics of pressure ulcers using photographs by staff nurses in acute care settings. *J Wound Ostomy Continence Nurs* 2013; 40: 150-6.