Assessing the dynamic changes in vitiligo: reliability and validity of the Vitiligo Disease Activity Score (VDAS) and Vitiligo Disease Improvement Score (VDIS)

N.van Geel,1, * L. Depaepe,2 V. Vandebe,1 L. Mertens,1 J. Van Causenbroeck,3 S. De Schepper,1 L. Van Coile,1 A. Van Reempts,1 A.-S. De Vos,1 J. Papeleu,1 I. Hoorens,1 D. Mertens,4 A. Wolkerstorfer,5 J.E. Lommerts,5 R. Speeckaert1
1Department of Dermatology, Ghent University Hospital, Ghent, Belgium
2Department of Dermatology, AZ Delta, Torhout, Belgium
3Department of Dermatology, Clinic Dr. med. Rotterdam, Gelsenkirchen, Germany
4Department of Dermatology, Universitary Clinic of Essen, Essen, Germany
5Department of Dermatology, Institute for pigment disorders and Infection & Immunity Institute Amsterdam UMC, Amsterdam, The Netherlands
*Correspondence: N. van Geel. E-mail: nanja.vangeel@ugent.be

Abstract

Background The assessment of the individual evolution of vitiligo is important for therapeutic decision making in daily practice. A fast, simple and validated physician-reported score to assess clinical changes in depigmentation over time in separate parts (activity and improvement) is currently missing.

Objective The main objective of the study was to develop and validate the Vitiligo Disease Activity Score (VDAS) and Vitiligo Disease Improvement Score (VDIS).

Methods The Vitiligo Disease Activity Score (VDAS) and Vitiligo Disease Improvement Score (VDIS) were evaluated based on a photo set of 66 patients with two different time points. In the first (short) version, only the number of changing body regions was counted based on 15 predefined areas (VDAS15 and VDIS15), while in the second (extensive) version the degree of worsening or improvement from \([-4\) to \(+4\) for each body area was added for a more detailed assessment (VDAS60 and VDIS60). Content and construct validity were tested. In addition inter-, intrarater reliability and feasibility were evaluated by 7 (test) and 5 (retest) physicians.

Results Evidence for content and construct validity was provided. Overall, VDAS15, VDIS15, VDAS60 and VDIS60 demonstrated good to excellent inter-rater reliability [intraclass correlation (ICC): VDAS: range = 0.797–0.900; VDIS: range = 0.726–0.798]. The intrarater reliability ICCs were 0.865 and 0.781 for the VDAS15 and VDIS15, respectively. Similar results were obtained for the VDAS60 and VDIS60 (ICC = 0.913 and 0.800, respectively). Completion time was short (median: 122 s/patient (first round); 95 s/patient (second round)).

Limitations Single tertiary centre mainly of skin phototype 2 to 3.

Conclusion The VDAS and VDIS appear to be valid, reliable and feasible instruments to score the evolution of vitiligo lesions. This accommodates the current urgent need for a simple, standardized and practical assessment of vitiligo activity and improvement over time.

Introduction

Vitiligo is a common (prevalence of 1–2%), acquired skin disease characterized by sharply demarcated depigmented lesions localized on any part of the body.1 The unpredictable and variable disease course, including periods of disease stability and activity, seriously impacts the quality of life of the patient.2 In clinical practice, the management of vitiligo is challenging and primarily based on the disease extent and the disease activity.3,4 This assessment should preferably be based on validated instruments. With regard to disease extent, the VASI is a commonly used tool. More recently the 'Vitiligo Extent Score’ (VES) and VESplus are developed as fast and accurate scoring instruments to measure the affected Body Surface Area (BSA) of patients with vitiligo.5–7 In addition to disease extent, the evaluation of disease activity is at least equally important.8 International guidelines provide specific interventions for patients with active vitiligo, while other interventions are indicated in stable vitiligo only.1,9 However, there is currently no agreed consensus on the grading of disease activity or on preferred instrument(s) for
measuring the degree of disease activity. This is especially important in the light of the development of new immunomodulating topical and systemic treatments for vitiligo.10,11

To date, disease activity is often assessed by asking the patient about recent progression of vitiligo. However, there is an urgent need for a better and standardized way to determine disease activity, as the initiation and choice of vitiligo treatment is highly depending on this. Disease activity can be assessed as disease progression over time (i.e. ‘dynamic’ assessment) or at a single time point (i.e. ‘static’ assessment) by examining visible clinical signs associated with disease activity. For the latter, we introduced recently the Vitiligo Signs of Activity Score (VSAS).12 It is based on the assessment of three clinical visible signs in vitiligo (e.g. confetti-like depigmentations, Koebner phenomenon and hypochromic areas/borders).13-15 To assess vitiligo activity over time, a comparison between two time points (‘dynamic’ assessment) with the aid of clinical pictures is most commonly used. In this study, we developed and validated the Vitiligo Disease Activity Score (VDAS) and Vitiligo Disease Improvement Score (VDIS) to assess the progression (activity) and repigmentation (improvement) in a standardized way. The combination of this new score with the VSAS will provide a complete set of information related to the quantification of the disease activity signs (‘static’ assessment) as well as changes over time (‘dynamic’ assessment).

**Materials and methods**

**Study design, ethics and construction of the scoring system**

This study was performed at the department of dermatology, Ghent University Hospital. During a preparatory phase, the scoring system was developed, evaluated and modified based on experience in clinical practice. Key factors kept in mind during development were relevance for patients, feasibility and usefulness both in clinical practice and in trials. Two pilot sessions were performed including 7 and 5 raters, respectively. Based on these pilot sessions some modifications were included resulting in a first (short) and a second (more extensive) version of the instrument. To score the course of the disease between two time points, a paper version of the instruments was used (see Fig. 1).

The study has been approved by the ethics committee (reference number Ghent: B670201421409). The COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist was used as a guide for designing and reporting our study.

**Raters and participants**

All raters were clinicians at the dermatology department of the Ghent University Hospital with different levels of experience, including dermatology residents (5), dermatologists (1) and a vitiligo expert (1). 3 of 7 raters were involved in the preceding pilot sessions using the same 66 patients. However, to avoid any recall bias between the pilot sessions and the final scoring rounds, an interval of many months (>3) was included between the scoring rounds.

Patients who were visiting our clinic and who provided written informed consent in the period between September 2018 and September 2019 were selected. Patients of all ages, with and without treatment were included, with the clinical diagnosis of non-segmental vitiligo and having at least two clinical photo sets with 6 (±2) or 12 (±2) months of interval. Patients with segmental vitiligo were excluded.

**Selection of photographs**

For each patient, two clinical photo sets were selected with 4–14 months of interval. In those who had multiple photo sets meeting this criterion, one of both intervals [6 (±2) or 12 (±2) months] was randomly assigned in order to obtain an equal number of patients for each interval. The selected photo sets were placed on Microsoft Powerpoint slides with each slide containing two photographs of one body area affected by vitiligo on two different dates. Only body areas present in both photo sets were used. Photographs were taken in a standardized manner during clinical practice and could consist of both UV and non-UV images.

**Scoring sessions**

The sequence of the slideshows was randomly picked (Research Randomizer: https://www.randomizer.org/) for repeating scoring rounds. All assessors received a training session of at least 15 min before the start of the first scoring round and a document with written instructions was provided.

**Reliability**

To check the reliability of the scoring system, the inter- and intrarater reliability were assessed. The scoring rounds were repeated by the raters with an interval of at least 2 weeks.

**Validity**

Evidence for content validity for the items included in the VDAS and VDIS was provided by the opinion of vitiligo experts based on a VGICC workshop in Rome 2016 (Nov 30-Dec 3rd).16 A questionnaire was filled out individually by 28 vitiligo experts. The question included in the questionnaire was referring to what items one suggests to include in a disease activity instrument based on a ‘dynamic’ assessment (difference between two time points). A five-point scale was used to provide an answer. For the interpretation of the results per item, a ‘consensus’ was defined as ‘if at least 70% “agreement” or “strong agreement” was reached’.

As no gold standard exists to evaluate criterion validity, the evaluation of construct validity was performed by testing against hypotheses. A draft of the hypotheses was formulated by two investigators (NvG and RS) both involved in the preceding pilot studies and partial analyses of this trial. To avoid a possible bias,
the draft of the hypotheses was subsequently evaluated by an external investigator (AW) not familiar with any previous data or results, and the magnitude of the relations was set. This investigator consulted a second external investigator (AL) for the final approval. Some of the hypotheses were based on a comparison (correlation) between the median VDAS (15 or 60) or VDIS (15 or 60) of all raters and an expert Physician’s Global Assessment (PGA) score for clinical disease evolutions (performed by one expert, who was not involved in the scoring sessions). This PGA score included a five-point scale ranging from no disease activity/no improvement to very severe disease activity/very much improved. Sufficient evidence for construct validity was assumed if ≥75% of the hypotheses were in agreement with the results.

**Feasibility**

The completion time per patient was recorded by each rater to evaluate the feasibility of this scoring system. After completion of the scoring rounds, all raters evaluated the user-friendliness on a five-point scale.

**Statistics and data analysis**

All statistical analyses were performed using IBM SPSS Statistics version 26 and Medcalc 19.8. For interrater reliability, a two-way random, absolute agreement, single-measures intraclass correlation coefficient (ICC) was used, and for intrarater reliability a two-way mixed, absolute agreement, single-measures ICC. For both inter-rater and intrarater reliability, an ICC of 0.75 or more was considered as excellent, between 0.6 and 0.74 as good, between 0.4 and 0.6 as fair and lower than 0.4 as poor. To test construct validity, Spearman’s rho correlation was used. To compare average completion times between the first and second rounds, a Wilcoxon signed-rank test was used.

**Results**

**Development of the VDAS and VDIS**

The scoring tools were constructed using similar designs for scoring activity (VDAS) and improvement (VDIS) (Fig. 1). The
assessor has to identify all the involved body areas in the first figure (0–15) and then indicate in the second and third figures (follow-up figures) if this vitiligo area shows activity (progression) and/or improvement (repigmentation) (= VDAS15 and VDIS15, respectively). The main numeric outputs obtained are the overall VDAS and VDIS values, ranging from 0 to 15. An additional option was added in each area showing the grade of difference, ranging from −4 to +4 (very much worsened or improved), resulting in a score of 0–60 (VDAS60 and VDIS60).

Raters and participants
The first scoring round was performed by seven raters and the retest by five raters. A total of 66 patients were included for the scoring rounds [Fitzpatrick skin type II: 14 (21.2%), III: 39 (59.1%), IV: 2 (3.0%), V: 1 (1.5%) and VI: 1 (1.5%), unknown: 9 (13.6%); female/male: 57.6%/42.4%]. The interval between retest by five raters. A total of 66 patients were included for the scoring rounds [Fitzpatrick skin type II: 14 (21.2%), III: 39 (59.1%), IV: 2 (3.0%), V: 1 (1.5%) and VI: 1 (1.5%), unknown: 9 (13.6%); female/male: 57.6%/42.4%]. The interval between treatments and score calculation was analysed separately (assessed by VESplus) varied between the patients [range: 0.01% – 36.43%, median (mean) VES: 1.3% (3.2%)]. The mean age at the relative percentage of worsening between 2 time points [rho (range) = 0.789; 95% CI: 0.730–0.856] and 0.861 (0.806–0.905); VDIS15: ICC = 0.726 (95% CI: 0.645–0.801) and 0.793 (95% CI: 0.722–0.855), respectively]. Inter-rater agreement for grading the changes within the involved areas (grade 0 to 4:) were in the first and second round good to excellent for VDAS60 (ICC = 0.856; 95% CI 0.804–0.901 and ICC = 0.900; 95% CI 0.857–0.933, respectively) and for VDIS60 (ICC = 0.766; 95% CI = 0.692–0.833 and ICC = 0.798; 95% CI: 0.724–0.860, respectively).

Validity
a) Content validity Content validity for the items included in the VDAS and VDIS was provided based on the results obtained during the VGICC workshop with vitiligo experts.16 Agreement on the items to include in a dynamic assessment of vitiligo disease activity were ‘Evolution in disease extent’ (e.g. worsening-stable-improved) (agree–strongly agree: 96%), ‘Number of active body locations’ (agree–strongly agree: 88%) and ‘Time interval of the 2 different time points’ (agree–strongly agree: 85%). As a possible indirect item ‘disease activity index’ was also scored but did not reach >70% agreement: ‘Number of active body locations/total number of affected body locations’ (agree–strongly agree: 65%).

b) Construct validity All hypotheses with the corresponding results can be found in Tables 1 and 2. For VDAS15, VDAS60, VDIS15 and VDIS60 sufficient evidence for construct validity was confirmed. Very strong correlations with the PGA expert were found for all scores [rho (range) = 0.757–0.864]. The VDAS15 and VDAS60 correlated strongly with the relative percentage of worsening between 2 time points as measured by the VESplus score (rho = 0.791 and rho = 0.795, respectively) while moderate correlations were found with the improvement (repigmentation) scores (VDIS15: rho = 0.372 and VDIS60: rho = 0.486). Patients with a longer follow-up interval (12 ± 2 months vs. 6 ± 2 months) and more involved body locations had higher VDAS15 and VDAS60 scores.

Reliability
a) Inter-rater reliability The median (mean) total scores (first round) were 6 (6; range 1–14, IQR 3–8) for the areas of involvement, 1 (1.56; range 0–9; IQR: 0–2) for areas with activity/worsening (VDAS15) and 1 (1.15; range 0–7; IQR: 0–1) for areas with improvement/repigmentation (VDIS15), respectively. The median total scores for grading the changes was 0.75 (2.44; range 0–23; IQR: 0–4) for activity/worsening (VDAS60) and 1 (1.71; range 0–10.; IQR: 0–2.5) for improvement/repigmentation (VDIS60) (Fig. 2). The inter-rater reliability for the number of involved areas was excellent (first round: ICC = 0.975; 95% CI = 0.965–0.983; second round: ICC = 0.971; 95% CI = 0.958–0.981). The inter-rater reliabilities for the VDAS15 and VDIS15 were good to excellent for both the first and second round [VDAS15: ICC = 0.797 (95% CI = 0.730–0.856) and 0.861 (0.806–0.905); VDIS15: ICC = 0.726 (95% CI: 0.645–0.801) and 0.793 (95% CI: 0.722–0.855), respectively]. Inter-rater agreement for grading the changes within the involved areas (grade 0 to 4:) were in the first and second round good to excellent for VDAS60 (ICC = 0.856; 95% CI 0.804–0.901 and ICC = 0.900; 95% CI 0.857–0.933, respectively) and for VDIS60 (ICC = 0.766; 95% CI = 0.692–0.833 and ICC = 0.798; 95% CI: 0.724–0.860, respectively).

Feasibility
Completion time In the first scoring round, the total completion time per patient which includes both filling out the scoring instruments (involved areas, VDAS and VDIS) and the degree of changes as well as the calculation of the score, was 122.4 (135.9) seconds/patient [median (mean)] (range 36.6–344.6). In the second round, this was 94.6 (92.5) seconds/patient [median (mean)] (range 25.8–192.6). A Wilcoxon signed-rank test showed a significant difference between the median completion times of both rounds (P < 0.001).

In addition, the median (mean) completion of the instruments and score calculation was analysed separately (assessed by four raters) and was 34.0 (36.0) seconds/patient (range: 8.25–92.5) for filling out the instrument (15 anatomic areas) and 73.5 (88.97) seconds/patient (range: 17.3–272) to calculate the scores, respectively (first round).
Table 1 Hypotheses for construct validity concerning the Vitiligo Disease Activity Score 0–15 (VDAS_{15}) and Vitiligo Disease Improvement Score 0–15 (VDIS_{15})

| Hypothesis                                                                 | Result                        | Confirmed (C) or failed (F) |
|----------------------------------------------------------------------------|-------------------------------|-----------------------------|
| **VDAS_{15}**                                                              |                               |                             |
| 1. We expect a rank correlation coefficient of at least 0.5 between the overall* VDAS_{15} and the PGA expert global disease progression score | Rho = 0.831 (95% CI: 0.734–0.895) | C                           |
| 2. We expect a rank correlation coefficient of at least 0.3 between the overall* VDAS_{0,15} and the relative percentage worsening (in %) between 2 time points of the VESplus score | Rho = 0.791 (95% CI: 0.675–0.869) | C                           |
| 3. Patients with a follow-up (FU) interval of 12 (± 2) months will have an at least 5% mean higher* VDAS_{15} than patients with an interval of 6 (± 2) months | FU 6 (± 2) months: VDAS_{15} = 0.794 – 2.375199% higher | C                           |
| 4. We expect that patients with more than 5 involved body locations* will have an at least 10% higher VDAS_{15} compared to patients with less than 6 involved body locations | N ≤ 5: VDAS_{15} = 0.625 N > 5: VDAS_{15} = 2.4412291% higher | C                           |

**VDIS_{15}**

1. We expect a rank correlation coefficient of at least 0.5 between the overall* VDIS_{15} and the PGA expert repigmentation score

2. We expect a rank correlation coefficient of at least 0.3 between the overall* VDIS_{0,15} and the relative percentage improvement (in %) between 2 time points of the VESplus score

3. In patients with an overall* VDIS_{0,15} of ≤6/15, the answer of the PGA expert global disease repigmentation scores will be at least ‘slightly improved (or more)’ in at least 50% of cases.

4. We expect that the mean repigmentation score on the face [(total sum score* for face/number of cases with involvement of the face)] will be ≥20% higher compared to the mean repigmentation score on the hands [(total sum score* for hands/number of cases with involvement of the hands)]

| Hypothesis | Result | Confirmed (C) or failed (F) |
|------------|--------|-----------------------------|
| **VDIS_{0,15}** |         |                             |
| 1. We expect a rank correlation coefficient of at least 0.5 between the overall* VDIS_{0,15} and the PGA expert global disease progression score | Rho = 0.864 (95% CI: 0.784–0.916) | C |
| 2. We expect a rank correlation coefficient of at least 0.3 between the overall* VDIS_{0,15} and the relative percentage worsening (in %) between 2 time points of the VESplus score | Rho = 0.795 (95% CI: 0.680–0.871) | C |
| 3. Patients with a follow-up (FU) interval of 12 (± 2) months will have an at least 5% mean higher* VDIS_{0,15} than patients with an interval of 6 (± 2) months | 211% higher | C |
| 4. We expect that patients with more than 5 involved body locations* will have an at least 10% higher VDIS_{0,15} compared to patients with less than 6 involved body locations | 303% higher | C |

**VDIS_{0}**

1. We expect a rank correlation coefficient of at least 0.5 between the overall* VDIS_{0} and the PGA expert repigmentation score

2. We expect a rank correlation coefficient of at least 0.3 between the overall* VDIS_{0,15} and the relative percentage improvement (in %) between 2 time points of the VESplus score

3. In patients with an overall* VDIS_{0} of ≤5/60, the answer of the PGA expert global disease repigmentation scores will be at least ‘slightly improved (or more)’ in at least 50% of cases

4. We expect that the mean repigmentation score on the face [(total sum score* for face/number of cases with involvement of the face)] will be ≥20% higher compared to the mean repigmentation score on the hands [(total sum score* for hands/number of cases with involvement of the hands)]

| Hypothesis | Result | Confirmed (C) or failed (F) |
|------------|--------|-----------------------------|
| **VDIS_{0}** |         |                             |
| 1. We expect a rank correlation coefficient of at least 0.5 between the overall* VDIS_{0} and the PGA expert repigmentation score | Rho = 0.822 (95% CI: 0.721–0.889) | C |
| 2. We expect a rank correlation coefficient of at least 0.3 between the overall* VDIS_{0,15} and the relative percentage improvement (in %) between 2 time points of the VESplus score | Rho = 0.486 (95% CI: 0.270–0.656) | C |
| 3. In patients with an overall* VDIS_{0} of ≤5/60, the answer of the PGA expert global disease repigmentation scores will be at least ‘slightly improved (or more)’ in at least 50% of cases | 100% of cases | C |
| 4. We expect that the mean repigmentation score on the face [(total sum score* for face/number of cases with involvement of the face)] will be ≥20% higher compared to the mean repigmentation score on the hands (total sum score* for hands/number of cases with involvement of the hands) | 93% higher | C |

* Median of all raters.
**User-friendliness** The median (mean) user-friendliness (scale 0–10) was 8.5 (8.5) for the VDAS15 and 8.5 (8.5) for the VDIS15. The user-friendliness for the separate additional grading scale +4 to -4 (one scale for both VDAS60 and VDIS60) was slightly lower but still excellent [= 8 (7.7)].

**Discussion**

Based on this study, the VDAS and VDIS appear to be valid, reliable and user-friendly measurement instruments to score the dynamic changes in depigmentation of vitiligo. This accommodates the current urgent need for a simple, standardized and practical assessment of activity and improvement over time. The large majority of dermatologists and patients agreed that cessation of spread should be part of a core domain set for vitiligo. It has been demonstrated that cessation of spread (= disease stability) is a valuable treatment goal for patients and should therefore be measured accordingly. For this, we developed a simple instrument (VDAS) to measure disease activity over time. Similar to other disorders, it seems logic to measure disease activity in vitiligo on a single continuous scale ranging from ‘worse’ to ‘improved’. However, areas with activity (worsening) and improvement (repigmentation) can occur in the same patient during follow-up, sometimes even in the same body area. In our cohort, 19/66 (28.8%) of patients displayed areas of repigmentation while other lesions progressed. As such, disease activity is not fully reflected by measuring the difference in disease extent between two time points. Valuable information concerning the dynamic changes in vitiligo patients would be lost without separate assessments of improvement (repigmentation) and activity (worsening). Different from other skin diseases, the improvement (repigmentation) of vitiligo not only needs a favourable immune environment but requires also additional stimulation of (precursor) melanocytes, usually by ultraviolet light. Additionally, as most treatments exert only an immunomodulating effect, their efficacy is not well reflected by the amount of repigmentation. Therefore, the VDAS and VDIS were developed and validated separately allowing to monitor the different aspects of disease evolution and the efficacy of treatments reliably.

The content validity, defined by an international group of experts, showed that an evolution score should contain...

---

**Figure 2** Vitiligo Disease Activity Score [VDAS15 (A), VDAS60 (C)] and Vitiligo Disease Improvement Score [VDIS15 (A) and VDIS60 (C)] values according to the median scores.
information on the number of changing locations. Additionally, as the efficacy of treatments depends strongly on the body location, the identification of which areas are changing is crucial. The VDAS and VDIS allow for an easy and rapid evaluation of the overall change while retaining specific information on each body location which is ideally suited for clinical practice. With new drugs for vitiligo on the horizon, it is essential to have a comprehensive set of validated measurement instruments to assess the extent and evolution of vitiligo lesions both in trials and clinical practice.

A simple and validated scoring tool to measure the multi-dimensional changes (worsening and improvement) in vitiligo was missing. The VDAS and VDIS are supported by a scoring sheet including a visual representation of the areas to be scored, which will allow a more standardized assessment. This further supports the comprehensibility and comprehensiveness of the tool. The final score as well as the scores per region can easily be implemented in any medical record (see example Fig. 1).

For each instrument, we validated two versions varying in items and scale dimensions (from 0 to 15 or 0 to 60). The first (VDAS15 and VDIS15) involves the total number of body locations with changes (activity or improvement) and can be useful for clinical practice due to its simplicity. The second version (VDAS60 and VDIS60) includes the option to grade the degree of activity/improvement per area and can be more interesting for clinical trials. The outcome of the 0–15 scale (= sum of the changing body areas) has the advantage of a direct clinical meaning while the option of grading the magnitude of change (VDAS60 and VDIS60) offers more details on the degree of improvement or worsening. For all scores, a good-to-excellent inter- and intrarater reliability was found. Construct validity was successfully tested by confirming all the predefined hypotheses.

Currently, the Vitiligo Disease Activity (VIDA) Scale, albeit in its original as a patient-reported outcome measure or a modified patient or physician-reported version, is a frequently used instrument that maps the last time when activity (worsening) of the vitiligo lesions was observed. However, the VIDA offers no information on the magnitude of the change. Moreover, the assessment of the disease evolution by patients can have limitations in vitiligo due to recall bias, slow changes over time and differences in skin tone depending on sun exposure in fair-skin

**Figure 3** Bland–Altman plots of the intrarater of score difference between the first and second round per rater plotted against the average score between both rounds per rater for the Vitiligo Disease Activity Score [VDAS15 (a); VDAS60 (c) and Vitiligo Disease Improvement Score (VDIS15 (b); VDIS60 (d)).
patients. The potential value of the assessment of disease activity by patients in clinical practice still needs to be further investigated and confirmed.

A limitation of this study is the low number of patients with dark skin types. Moreover, photographic material used in this study was suboptimal which could have negatively affected the results for the reliability testing. In addition, as this study was conducted in one centre, the observed changes might not be generalizable to the global population of vitiligo patients.

In conclusion, this study introduces and validates measurement instruments for both activity and improvement in vitiligo. These scores can be used both in clinical practice as in trials combining good reliability, content validity, construct validity, user-friendlyness and favourable timing.

Acknowledgement
We would like to express our gratitude to the volunteering patients for the use of their pictures during the scoring rounds.

Funding sources
This project was supported by a grant from Incyte Biosciences International Sarl. The opinions expressed in this paper are those of the authors and do not represent those of Incyte Biosciences. The research activities of N. van Geel and R. Speeckaert are supported by the Scientific Research Foundation-Flanders (FWO Senior Clinical Investigator: 1831512N (NvG) and 18B2721N (RS), respectively).

Conflict of interest
NvG: Consultancy and/or investigator: Pfizer, Incyte, Sunpharma and Abbvie. NvG was involved in the preparative phase (e.g. design and pilot testing) of the scores used. The co-authors Virginie Vandaele, Laura Mertens, Jérôme Van Causenbroeck, Sofie De Scheepere, Laura Van Coile, Astrid Van Reempts, Ann-Sophie De Vos, Jorien Papeleu, Isabelle Hoorens, Delphine Mertens, Albert Wolkerstorfer, Janny E Lommerts, Reinhart Speeckaert declare no conflict of interest related to the content of the paper.

Data availability statement
The design, material and methods that support the findings of this study are available from the corresponding author upon request. The database is not available due to privacy or ethical restrictions.

References
1 Ezzeidine K, Eleftheriadou V, Whitton M, van Geel N. Vitiligo. Lancet 2015; 386: 74–84.
2 Linthorst Homan MW, Spuls PI, de Korte J, Bos JD, Sprangers MA, van der Veen JPW. The burden of vitiligo: patient characteristics associated with quality of life. J Am Acad Dermatol 2009; 61: 411–420.
3 Speeckaert R, van Geel N. Vitiligo: an update on pathophysiology and treatment options. Am J Clin Dermatol 2017; 18: 733–744.
4 Eleftheriadou V, Atkar R, Batchelor J et al. British Association of Dermatologists guidelines for the management of people with vitiligo 2021. Br J Dermatol 2021; 186: 18–29.
5 van Geel N, Lommerts J, Bekkenk M et al. Development and validation of the Vitiligo extent score (VES): an international collaborative initiative. J Invest Dermatol 2016; 136: 978–984.
6 van Geel N, Lommerts JE, Bekkenk MW et al. Development and validation of a patient-reported outcome measure in vitiligo: the self assessment Vitiligo extent score (SA-VES). J Am Acad Dermatol 2017; 76: 464–471.
7 van Geel N, Wolkerstorfer A, Lommerts JE et al. Validation study of the Vitiligo extent score-plus. J Am Acad Dermatol 2018; 78: 1013–1015.
8 van Geel N, Desmedt V, De Schepper S, Boone B, Laperehe H, Speeckaert R. Cessation of spread as a treatment objective in vitiligo: perception from the patients’ point of view. Br J Dermatol 2016; 174: 922–924.
9 Taieb A, Alomar A, Böhm M et al. Guidelines for the management of vitiligo: the European dermatology forum consensus. Br J Dermatol 2013; 168: 5–19.
10 Rosmarin D, Pandya AG, Lebowohl M et al. Ruxolitinib cream for treatment of vitiligo: a randomised, controlled, phase 2 trial. Lancet 2020; 396: 110–120.
11 Liu LY, Strassner JP, Refta MA, Harris JE, King BA. Repigmentation in vitiligo using the janus kinase inhibitor, tofacitinib, may require concomitant light exposure. J Am Acad Dermatol 2017; 77: 675–682.e1.
12 van Geel N, Passeron T, Wolkerstorfer A, Speeckaert R, Ezzeidine K. Reliability and validity of the Vitiligo signs of activity score (VSAS). Br J Dermatol 2020; 183: 883–890.
13 Zhang L, Chen S, Kang Y et al. Association of Clinical Markers with Disease Progression in patients with Vitiligo from China. JAMA Dermatol 2020; 156: 288–295.
14 Sosa JJ, Currimbhoy SD, Ukoha U et al. Confetti-like depigmentation: a potential sign of rapidly progressing vitiligo. J Am Acad Dermatol 2015; 73: 272–275.
15 van Geel N, Grine L, De Wispelaere P, Mertens D, Prinsen CAC, Speeckaert R. Clinical visible signs of disease activity in vitiligo: a systematic review and meta-analysis. J Eur Acad Dermatol Venereol 2019; 33: 1667–1675. https://doi.org/10.1111/jdv.15604.
16 van Geel N, Boniface K, Senneschal J et al. Meeting report: Vitiligo global issues consensus conference workshop “outcome measurement instruments” and vitiligo international symposium, Rome, Nov 30-Dec 3rd. Pigment Cell Melanoma Res 2017; 30: 436–443.
17 Eleftheriadou V, Thomas K, van Geel N et al. Developing core outcome set for vitiligo clinical trials: international e-Delphi consensus. Pigment Cell Melanoma Res 2015; 28: 363–369.
18 Thomas KS, Batchelor JM, Akram P et al. Randomized controlled trial of vitiligo topical corticosteroid and home-based narrowband ultraviolet B for active and limited vitiligo: results of the HI-light Vitiligo trial. Br J Dermatol 2021; 184: 828–839.
19 Mehta H, Kumar S, Parsad D, Bishnoi A, Vinay K, Kumanar MS. Oral cyclosporine is effective in stabilizing active vitiligo: results of a randomized controlled trial. Dermatol Ther 2021; 34: e15033. https://doi.org/10.1111/dth.15033.
20 Brazzelli V, Antoninetti M, Palazzini S, Barbagallo T, De Silvestri A, Borroni G. Critical evaluation of the variants influencing the clinical response of vitiligo: study of 60 cases treated with ultraviolet B narrow-band phototherapy. J Eur Acad Dermatol Venereol 2007; 21: 1369–1374.
21 Wolkerstorfer A. The long road to valid outcomes in vitiligo. Br J Dermatol 2019; 180: 454–455.
22 Njoo MD, Das PK, Bos JD, Westerhof W. Association of the Körner phenomenon with disease activity and therapeutic responsiveness in vitiligo vulgaris. Arch Dermatol 1999; 135: 407–413.
23 Coias J, Hynan LS, Pandya AG. Lack of correlation of the patient-derived Vitiligo disease activity index with the clinician-derived Vitiligo area scoring index. J Am Acad Dermatol 2018; 78: 1013–1016.