Validation of the Vitiligo Noticeability Scale: a patient-reported outcome measure of vitiligo treatment success*

J.M. Batchelor,[1] W. Tan,[2] S. Tour,[1] A. Yong,[3] A.A. Montgomery[2] and K.S. Thomas[1]

[1]Centre of Evidence Based Dermatology, King's Meadow Campus, Lenton Lane, Nottingham NG7 2NR, U.K.
[2]Nottingham Clinical Trials Unit, Queen's Medical Centre, C Floor, South Block, Nottingham NG7 2UH, U.K.
[3]Norfolk and Norwich University Hospitals NHS Foundation Trust, Norwich NR4 7UY, U.K.

Summary

Background Patient-reported outcome measures are rarely used in vitiligo trials. The Vitiligo Noticeability Scale (VNS) is a new patient-reported outcome measure assessing how 'noticeable' vitiligo patches are after treatment. The noticeability of vitiligo after treatment is an important indicator of treatment success from the patient's perspective.

Objectives To evaluate the construct validity, acceptability and interpretability of the VNS.

Methods Clinicians (n = 33) and patients with vitiligo (n = 101) examined 39 image pairs, each depicting a vitiligo lesion pre- and post-treatment. Using an online questionnaire, respondents gave a global assessment of treatment success and a VNS score for treatment response. Clinicians also estimated percentage repigmentation of lesions (< 25%; 25–50%; 51–75%; > 75%). Treatment success was defined as 'yes' on global assessment, a VNS score of 4 or 5, and > 75% repigmentation. Agreement between respondents and the different scales was assessed using kappa (κ) statistics.

Results Vitiligo Noticeability Scale scores were associated with both patient- and clinician-reported global treatment success ($\kappa = 0.54$ and $\kappa = 0.47$, respectively). Percentage repigmentation showed a weaker association with patient- and clinician-reported global treatment success ($\kappa = 0.39$ and $\kappa = 0.29$, respectively). VNS scores of 4 or 5 can be interpreted as representing treatment success. Images depicting post-treatment hyperpigmentation were less likely to be rated as successful.

Conclusions The VNS is a valid patient-reported measure of vitiligo treatment success. Further validation of the VNS is required, using larger sets of clinical pre- and post-treatment images, affecting a wider range of anatomical sites.

What's already known about this topic?

- The lack of standardized outcome measures makes comparison of vitiligo treatment efficacy difficult.
- Patient-reported outcome measures are rarely used in vitiligo trials.
- Patient-reported outcome measures assessing vitiligo treatment success from the patient's perspective have yet to be developed.
- The Vitiligo Noticeability Scale (VNS) is a new patient-reported outcome measure of treatment response, which has been shown to have face validity.

What does this study add?

- The VNS has good construct validity, acceptability and interpretability, supporting its inclusion as a patient-reported measure of the cosmetic acceptability of treatment response in vitiligo trials.

- The VNS is a better and more consistent indicator of global treatment success than percentage repigmentation.
- VNS scores of 4 or 5 can be interpreted as representing treatment success.
- Further validation of the VNS is required.

Vitiligo is a chronic depigmenting disease, with a prevalence of 0·5–1·0%. Various treatments have been evaluated, including topical and oral preparations, light therapy, surgical procedures, psychological and complementary therapies.[1,2]

Despite the large number of randomized controlled trials (RCTs) assessing vitiligo treatments, the lack of standardized outcome measures makes comparison of treatment efficacies difficult.[2–5] A systematic review reported 25 different outcome measures across 54 trials.[4] Patient-rated outcomes such as cosmetic acceptability were assessed in just nine of 54 (17%) of the trials. Although percentage repigmentation was reported in 96% of trials (usually with ≥ 75% repigmentation representing treatment success), it was measured in 48 different ways, making comparison impossible.[4] Moreover, there have been no attempts to validate the use of ≥ 75% repigmentation, assessed by a clinician, to represent treatment success. To enable treatment comparison, work has been done to identify a set of core outcome measures that can be used in RCTs. International consensus has been reached for seven outcome domains, including repigmentation and cosmetic acceptability of treatment response.[6]

Inclusion of patient-rated outcome measures in RCTs of vitiligo treatments has been recommended,[2,4,5] but validated measurement scales are limited. Scales to assess psychosocial impact and quality of life in patients with vitiligo have recently been validated.[7–9] These scales include some elements that relate to the appearance of the vitiligo, but they cannot be used specifically to assess the appearance of the vitiligo after treatment or perceived 'treatment success' from the patient's perspective.

The Vitiligo Noticeability Scale (VNS) is a patient-rated outcome measure of vitiligo treatment response. We previously developed this scale through a process using consensus methodology, involving 165 patients.[10] This involved an online survey, to identify which aspects of treatment response are most important to patients, and online discussion groups, to explore the survey results with patients, and reach consensus regarding the wording for a proposed patient-reported outcome measure. The phrase 'cosmetically acceptable response' was not meaningful to people with vitiligo, as it seemed vague and impersonal. In contrast, they thought that 'how noticeable' the vitiligo is, the 'colour match' between their vitiligo and normal skin and a 'reduction in the size of the vitiligo patches' were more meaningful measures of treatment response. The participants reached consensus that the 'noticeability' of the vitiligo was the most important of these concepts for assessing the success of treatment response; a scale with five response options (both words and numbers)

would be the best scale to use when assessing treatment response (whereas giving a binary 'yes'/'no' response was more difficult); and a score of 4 or 5 on the scale would represent a successful treatment response.[8] Table 1 shows the details of the VNS.

This article reports on work to further validate the VNS. The main hypothesis was that the VNS would be a better and more consistent indicator of treatment success than percentage repigmentation.

In relation to this main hypothesis, we aimed to answer the following questions: 'What is the construct validity of the VNS (compared with global treatment success as reference standard), and how does VNS compare with clinician-rated percentage repigmentation?'. Construct validity is defined as the degree to which scores of a given measurement instrument are consistent with hypotheses relating to the relationship between that measurement instrument and other relevant scales.[11] Construct validity is established by testing specific hypotheses.

In relation to the VNS we hypothesized that (i) there would be a positive association between VNS and global treatment success, with a kappa (κ) statistic of ≥ 0·4; (ii) the VNS would have better association with global treatment success than with percentage repigmentation; and (iii) clinician- and

Table 1 Questions used to assess pre- and post-treatment images

Question	Response options (score)	Success criteria
VNS		
Compared with before treatment, how noticeable is the vitiligo now?	More noticeable (1) As noticeable (2) Slightly less noticeable (3) A lot less noticeable (4) No longer noticeable (5)	Score of 4 or 5
Global treatment success		
Do you think this treatment was successful? If 'no' please state why	Yes No	Yes
Percentage repigmentation		
Please state the level of repigmentation you think has been achieved in the 'after treatment' image	< 25% 25–50% 51–75% > 75%	> 75%

VNS, Vitiligo Noticeability Scale.

patient-assessed VNS scores would show agreement, with $\kappa \geq 0.4$.

The second question was as follows: 'What is the acceptability and interpretability of the VNS?'. The *acceptability* of a measurement instrument refers to whether or not an individual is willing to give a response to it, and how easy it is for them to do so.[11] The present study included an assessment of how easy the VNS is to use.

The *interpretability* of a measurement instrument refers to 'the degree to which one can assign qualitative meaning – that is, clinical or commonly understood connotations – to an instrument's quantitative scores or change in scores'.[11] A score of ≥ 4 on the five-point VNS had previously been proposed to represent treatment success.[10] The present study sought to further investigate this proposed cut-off, comparing it with global treatment success, in order to assess interpretability of the VNS.

The third question was: 'What factors influence VNS scores?'. We sought to explore factors that may influence the VNS scores. These included characteristics of the vitiligo patches, characteristics of the patients and characteristics of the clinicians.

Materials and methods

This validation study was performed using an online questionnaire, completed by people with vitiligo and clinicians (dermatologists). Approval for the study was granted by the University of Nottingham's research ethics committee (Ref. LTg15082013 SoM Dermatol).

Participants and setting

People with vitiligo (n = 188) on a mailing list held at the Centre of Evidence Based Dermatology (CEBD), University of Nottingham, received an email invitation to participate. The list included participants from a Vitiligo Priority Setting Partnership,[3] or those who had expressed interest in being involved in vitiligo research. Participants were also recruited through the Facebook page of the U.K. Vitiligo Society, and the survey's web link was tweeted through the U.K. Dermatology Clinical Trials Network Twitter account (http://www.ukdctn.org).[12] The clinician questionnaire was sent to U.K.-based dermatologists who had previously expressed an interest in recruiting patients for a vitiligo RCT (n = 21), and dermatologists involved in the international initiative to develop core outcomes for vitiligo (n = 51).

Pre- and post-treatment digital images

We created a bank of images showing vitiligo lesions pre- and post-treatment.

We sourced 10 pretreatment images of vitiligo from (i) a database held at the CEBD (with written consent from patients); (ii) published papers (after obtaining relevant permissions); and (iii) images in the public domain. We chose images showing vitiligo on the face (n = 27) and hands

(n = 12) as 'high-expression sites', where vitiligo is more visible. To ensure broad demographic coverage, we selected images showing both sexes (male, n = 14; female, n = 14; neutral, n = 11) with a range of skin phototypes (Table S1; see Supporting Information).

For four of the 'pretreatment' images, a suitable paired 'post-treatment' image was available, showing the same patch of vitiligo after treatment. We created other image pairs via image manipulation: the pretreatment image was copied and digitally altered using Adobe® Photoshop® CS6 Extended (predominantly clone stamp, healing brush, spot healing and patch tools; Adobe, San Jose, CA, U.S.A.) to simulate varying degrees of skin repigmentation, in a 'post-treatment' image. In total, we created 35 image pairs, using the same pretreatment images to create sets of post-treatment images. We created more images to show approximately 70–80% repigmentation (determined by the clinician authors), as clinician-assessed repigmentation of $\geq 75\%$ is often taken to represent treatment success.[4] The 'post-treatment' images showed common patterns of repigmentation – perifollicular (n = 13), marginal (n = 13) and diffuse (n = 11) – with some images also showing hyperpigmentation (n = 6).[13] We created a total of 39 'post-treatment' images.

Data collection

We created a questionnaire using the online survey tool Survey Monkey (http://www.surveymonkey.com). We made one version of the questionnaire for people with vitiligo and another for clinicians. We asked three people with vitiligo and three clinicians to pilot test the questionnaires.

We sent potential patient participants an explanatory email, with a web link to the questionnaire. We sent email reminders on three occasions and placed repeated posts on the Vitiligo Society Facebook page. To encourage participation, we pledged a £10 donation to the Vitiligo Society for every completed questionnaire received.

The questionnaire included questions about the respondents' vitiligo (if applicable) and basic demographic characteristics, followed by the 39 pairs of pre- and post-treatment images, with each pair presented on a new screen. Respondents were asked to assess each image pair using the two questions shown in Table 2. Examples of 'pre- and post-' image pairs are shown in Figure 1. (Note: both of the post-treatment images are simulated images.) Clinicians were additionally asked to rate percentage repigmentation in the post-treatment image (< 25%, 25–50%, 51–75%, > 75%).

Image pairs were presented in a random sequence and respondents were permitted to navigate only forwards through the questionnaire; they could not edit previous responses.

To assess the acceptability of the VNS, after assessing all of the 39 image pairs, respondents were asked the question, 'How easy was it for you to judge the answers to the questions given for each set of images?' (using a five-point scale: very easy, easy, okay, difficult, very difficult). Respondents could also add their own comments.

Table 2 Demographic and clinical characteristics of respondents

	Patient respondents (n = 101)	Clinician respondents (n = 33)
Sex		
Male	35 (35)	17 (52)
Female	65 (64)	16 (48)
NS	1 (1)	0
Age (years)[a]		
< 16	10 (10)	0
17–30	14 (14)	1 (3)
31–45	24 (24)	11 (33)
46–65	36 (36)	19 (58)
> 65	15 (15)	1 (3)
NS	2 (1)	1 (3)
Duration of vitiligo		
6–12 months	1 (1)	–
1–2 years	5 (5)	–
2–5 years	11 (11)	–
5–10 years	13 (13)	–
> 10 years	71 (70)	–
Approximate percentage of skin affected		
< 10	24 (24)	–
10–25	38 (24)	–
25–50	17 (16)	–
50–80	15 (15)	–
> 80	7 (7)	–
Ethnicity		
White British	70 (69)	6 (18)
Other white	11 (11)	10 (30)
Asian/mixed	11 (11)	4 (12)
Black/mixed	5 (5)	3 (9)
Other	3 (3)	9 (27)
NS	1 (1)	1 (3)
Country of residence		
U.K.	79 (78)	14 (42)
Australia	4 (4)	0
Canada	2 (2)	0
Egypt	3 (3)	7 (21)
U.S.A.	5 (5)	2 (6)
Other[b]	8 (8)	10 (30)

Data are n (%). NS, not stated. [a]Eight respondents answered the survey on someone else's behalf. [b]Other countries included Brazil, China, Costa Rica, France, Greece, Ireland, Japan, Latvia, Malaysia and Norway.

Sample size

As all three of our hypotheses about the VNS were based on estimating and comparing κ statistics, we aimed to have sufficient precision for these statistics [defined as half of the width of the 95% confidence interval (CI) of κ]. We investigated the association between sample size and the precision of the κ estimate using simulations. We created multiple simulated data sets of varying size, in which each observation consisted of 39 binary ratings of treatment success or failure. To simulate our survey, these 39 scores were clustered as 13 sets of three. By varying the intracluster correlation coefficient for treatment response at this level, and at observation level, and the overall

proportion of successes, we found that for sample sizes > 100 the precision of κ reached a plateau. Therefore, we aimed to include at least 100 patient respondents in the survey.

Statistical analyses

We performed all analyses with STATA version 13 (Stata Corp., College Station, TX, U.S.A.). We included all available data in the analyses and did not impute any missing data.

There is no gold standard for assessment of treatment success in patients with vitiligo, so as a reference standard we used patient- and clinician-reported treatment success on a 'yes/no' binary scale (see Table 1). We estimated crude agreement between the scales and the reference standard by converting the five-point VNS and the four-point percentage repigmentation scales into binary measures of success (success = VNS scores ≥ 4; percentage repigmentation > 75%). We assessed the extent of agreement between these binary scales using κ statistics (two outcomes, more than one rater) separately for patients and clinicians, with bootstrapping to estimate 95% CIs,[13] and allowing for clustering at both the level of the respondent and image 'set' defined by the 10 pretreatment images. We used the repigmentation category chosen by the majority of the clinicians to assign each image to the relevant percentage repigmentation category.

We estimated agreement between patients and clinicians on the VNS by calculating separate κ values for all possible patient–clinician pairs and taking the mean, with the 95% CI estimated by bootstrapping.[14] The Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist states that for a scale to be acceptable, correlations between two instruments measuring the same construct should be $\kappa \geq 0.50$, or 75% of the results should be in accordance with *a priori* hypotheses.[15] However, in this study we took a $\kappa \geq 0.4$ to show an acceptable (moderate) level of agreement with global treatment success, as this unvalidated scale may assess a slightly different but related construct.

Factors influencing VNS scores were explored using multivariable logistic regression. Characteristics of the vitiligo patches included in the logistic regression model were anatomical location of the lesion, pattern of repigmentation, sex of the person in the image, level of repigmentation, and presence or absence of hyperpigmentation. We included patients' characteristics in the patient-rated VNS treatment success model and clinicians' characteristics in the clinician-rated VNS treatment success model. Characteristics of the respondents included in the logistic regression model were age, sex, ethnicity, country of residence, level of expertise (clinician respondents only), and extent and duration of the vitiligo (patient respondents only).

Results

The online questionnaire was open from December 2013 to March 2014. There were 101 complete responses to the patient questionnaire; eight were from proxy respondents. The
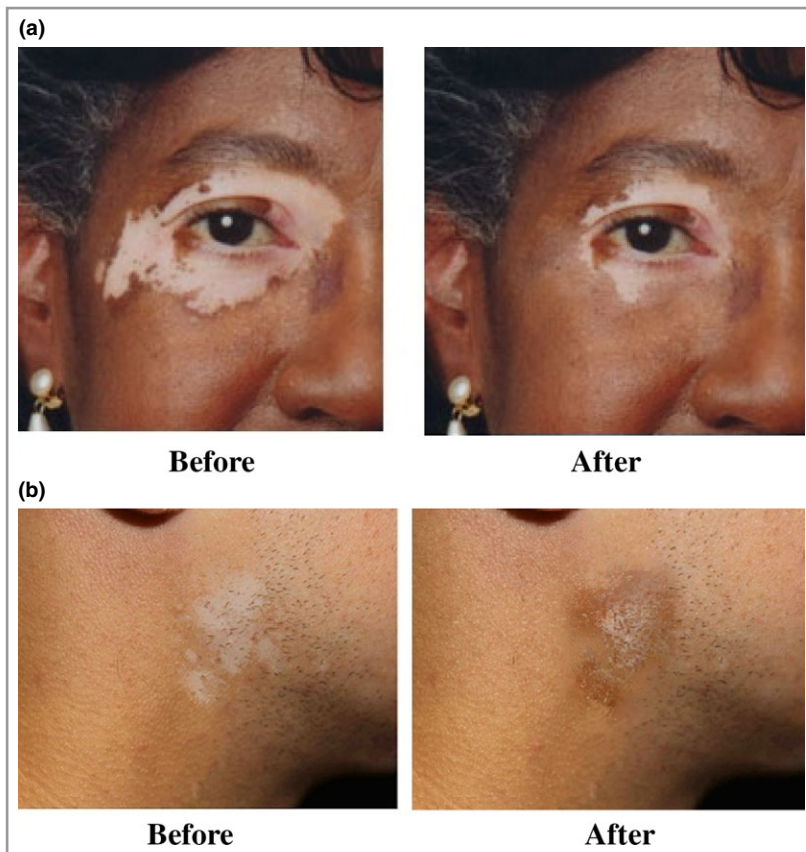
Fig 1. Examples of pre- and post-treatment image pairs (note: post-treatment images are simulated). (a) Female face, type V/VI skin, marginal repigmentation (approximately 50%). (b) Male face, type I/II skin, perifollicular repigmentation (approximately 70%) with hyperpigmentation.

clinician questionnaire was completed by 33 dermatologists, 18 (54%) of whom had a special interest in vitiligo. The characteristics of the respondents are shown in Table 2. While the majority of respondents were from the U.K., patients and clinicians from 15 other countries also took part. A small number of questions within the survey did not receive responses. For the 101 respondents to the patient questionnaire, 3613 out of a possible 3939 (92%) responses to the VNS and global success ratings were received. For the 33 clinicians, 1217 out of a possible 1287 responses (94%) were received for the three scales.

Question 1: 'What is the construct validity of the Vitiligo Noticeability Scale (compared with global treatment success as reference standard), and how does the Vitiligo Noticeability Scale compare with clinician-rated percentage repigmentation?'

We compared scores on the VNS and percentage repigmentation scale with global treatment success. Crude agreement in classifying treatment response using the VNS compared with global treatment success scale was 78% and 76% for patients and clinicians, respectively. Crude agreement in classifying treatment response using percentage repigmentation compared with global treatment success was 62% for clinicians (Table 3).

The VNS showed good-to-moderate agreement with patient- and clinician-reported global treatment success ($\kappa = 0.54$ and $\kappa = 0.47$, respectively). This supports our first hypothesis – that there would be a positive association between VNS and global treatment success, with $\kappa \geq 0.4$.

In contrast, percentage repigmentation showed only fair agreement with patient- and clinician-reported global treatment success ($\kappa = 0.39$ and $\kappa = 0.29$, respectively). This, together with the better levels of agreement between the VNS and global treatment success, supports the second hypothesis – that the VNS would have better association with global treatment success than percentage repigmentation.

Comparing patient- and clinician-rated VNS scores showed a moderate-to-good association ($\kappa = 0.43$). This supports our third hypothesis that clinician- and patient-assessed VNS scores would show agreement, with a $\kappa \geq 0.4$.

Question 2: 'What is the acceptability and interpretability of the Vitiligo Noticeability Scale?'

The majority of patients (57%) reported that the VNS was 'very easy' or 'easy' to complete, and 36% reported that it was 'okay'. Only 7% said it was 'difficult' to complete and none said it was 'very difficult'. These results suggest that the VNS is an acceptable and intuitive scale for completion by patients.

While the five-point VNS may be reported descriptively for all levels, the proposed cut-off for a successful treatment response was a score of ≥ 4 points.[10] The current study supported this classification: 94% and 99% of the images were

Table 3 Comparison of global treatment success vs. the Vitiligo Noticeability Scale (VNS) and percentage repigmentation

| | Patients (n = 101) | | Clinicians (n = 33) | |
| | Treatment success | | Treatment success | |
	No	Yes	No	Yes
VNS score[a]				
1 = more noticeable	**103 (95)**	5 (5)	**31 (84)**	6 (16)
2 = as noticeable	**532 (87)**	81 (13)	**110 (73)**	42 (27)
3 = slightly less noticeable	**342 (35)**	624 (65)	**109 (35)**	204 (65)
4 = a lot less noticeable	92 (6)	**1552 (94)**	37 (6)	**594 (94)**
5 = no longer noticeable	2 (1)	**280 (99)**	3 (4)	**81 (96)**
Crude agreement (%)[b]	**78**		**76**	
Percentage repigmentation[c]				
0–25			**109 (79)**	29 (21)
25–50			**90 (44)**	116 (56)
50–75			**58 (17)**	283 (83)
> 75			33 (6)	**499 (94)**
Crude agreement (%)[b]			**62[d]**	

Data are n (%) unless otherwise indicated. [a]Treatment success for VNS ≥ 4. [b]Crude agreement: figures in bold represent agreement between VNS or percentage repigmentation as binary variables and global treatment success. Estimate of crude agreement for patient-reported VNS: $(103 + 532 + 342 + 1552 + 280)/3613 = 78\%$. [c]Treatment success for percentage repigmentation > 75%. [d]Agreement within percentage repigmentation scale among clinicians: $\kappa = 0.26$.

classed as a treatment success for VNS scores of 4 and 5, respectively, and 95% and 87% of the images classed as unsuccessful for VNS scores of 1 and 2, respectively (Table 3). Nevertheless, a score of 3 on the VNS scale was considered to be a successful treatment for 65% of the image pairs, suggesting that a VNS score of 3 may represent treatment success for some patients.

Question 3: 'What factors influence Vitiligo Noticeability Scale scores?'

For patient respondents, the percentage and pattern of repigmentation, and presence of hyperpigmentation, were all associated with rating the treatment as successful (Table 4). The age and country of residence of patient respondents, and the age of clinician respondents, were also associated with rating the treatment as successful (Table 4).

The number of clinician respondents was relatively small and, compared with nondermatologists, all were likely to have some expertise in treating vitiligo, so the division between 'experts' and 'nonexperts' in this study is less clear than it would be between 'dermatologists' and 'nondermatologists'. Therefore, although we found no clear evidence in this study that the level of vitiligo expertise had any association with rating treatment as successful, it is not possible to draw firm conclusions from this dataset.

Discussion

Until now there have been no validated scales allowing people with vitiligo to rate treatment response. This study suggests that the patient-reported VNS has good construct validity,

acceptability and interpretability, supporting its inclusion as a patient-reported measure of the cosmetic acceptability of treatment response in vitiligo RCTs.

Patients with vitiligo were central to the development of the VNS,[10] and so it is perhaps unsurprising that VNS scores correspond well with patients' views of what constitutes treatment success. This study would suggest that a VNS score of 4 or 5 (i.e. 'a lot less noticeable' or 'no longer noticeable') should be interpreted as 'successful treatment response'; a VNS score of 3 as 'partially successful'; and a VNS score of 1 or 2 as 'unsuccessful'. This may be a better representation of the entire clinical picture than a binary division of scores into 'successful' (VNS score 4 or 5) and 'unsuccessful' (VNS score 1, 2 or 3).

The results shown here support our *a priori* hypotheses regarding the level of agreement expected for the different comparisons. This supports our main hypothesis that the VNS is a better and more consistent indicator of global treatment success than percentage repigmentation. Agreement was better between the patient-reported scales than between the clinician-reported scales, lending support to the use of VNS as a patient-rated measure.

The observation that hyperpigmentation of vitiligo patches has an adverse effect on patient-reported treatment success may explain the lower level of agreement between percentage repigmentation and treatment success. Percentage repigmentation measures the amount of repigmentation, regardless of the visual appearance of the skin, and so may overestimate treatment 'success'. Similarly, in cases where repigmentation is diffuse, there may be some restoration of pigment across the entire patch, which could be interpreted as '100% repigmentation'; however, the overall colour of the patch may still be

Table 4 Logistic regression of predictors of treatment success on the Vitiligo Noticeability Scale (VNS)

Variable	n	Unadjusted OR	95% CI	P-value	Adjusted OR	95% CI	P-value
Logistic regression of patient-rated VNS success on image characteristics							
Site							
Face	27	–	–		–	–	
Hands	12	0·77	0·18–3·34	0·72	0·77	0·26–2·28	0·63
Sex							
Female	14	–	–		–	–	
Male	13	1·88	0·47–7·46		1·54	0·83–2·85	
Unknown	12	1·04	0·20–5·47	0·61	–[a]	–	0·17
Pattern							
Marginal	14	–	–		–	–	
Perifollicular	13	1·07	0·22–5·13		0·73	0·31–1·71	
Diffuse	10	4·01	0·88–18·27		1·76	0·67–4·59	
Mixed	1	0·12	0·03–0·46		0·25	0·13–0·50	
Complete	1	8·86	2·34–33·56	< 0·01	4·85	1·26–18·68	< 0·01
Hyperpigmentation							
No	33	–	–		–	–	
Yes	6	0·40	0·09–1·69	0·21	0·10	0·06–0·18	< 0·10
Percentage repigmentation[b]	39	1·08	1·05–1·10	< 0·01	1·08	1·06–1·10	< 0·01
Logistic regression of patient-rated VNS success on patients' characteristics							
Response on behalf of							
Self	93	–	–		–	–	
Other	8	1·17	0·76–1·79	0·47	1·04	0·55–1·98	0·90
Sex							
Male	35	–	–		–	–	
Female	65	0·83	0·60–1·15	0·28	0·90	0·65–1·24	0·52
Age (years)							
< 16	10	–	–		–	–	
17–30	14	1·39	0·91–2·14		2·03	1·30–3·17	
31–45	24	1·01	0·65–1·57		1·30	0·84–2·01	
46–65	36	0·81	0·50–1·32		1·21	0·77–1·90	
> 65	15	0·86	0·56–1·33	0·07	1·44	0·93–2·22	0·03
Duration of vitiligo (years)							
< 2	61	–	–		–	–	
2–5	11	0·67	0·36–1·22		0·60	0·27–1·31	
5–10	13	0·53	0·31–0·92		0·58	0·27–1·23	
> 10	71	0·57	0·36–0·92	0·10	0·54	0·25–1·14	0·44
Percentage skin affected							
< 10	24	–	–		–	–	
10–25	38	0·83	0·56–1·24		0·81	0·56–1·17	
25–50	17	0·84	0·54–1·30		0·89	0·54–1·47	
50–80	15	0·84	0·48–1·44		0·86	0·52–1·42	
80	7	0·66	0·35–1·22	0·73	0·65	0·35–1·20	0·62
Ethnicity							
White	81	–	–		–	–	
Other	30	1·16	0·77–1·74	0·48	0·67	0·37–1·21	0·19
Country of residence							
U.K.	79	–	–		–	–	
Other	22	1·55	1·09–2·18	0·01	1·81	0·94–3·49	0·08
Logistic regression of clinician-rated VNS success on clinicians' characteristics							
Sex							
Male	17	–	–		–	–	
Female	16	0·94	0·66–1·33	0·72	1·00	0·73–1·36	1·00
Age (years)							
≤ 45	12	–	–		–	–	
> 45	21	1·45	1·06–1·96	0·02	1·44	1·05–1·97	0·02
Ethnicity							
White	16	–	–		–	–	
Other	17	0·83	0·59–1·18	0·30	0·92	0·68–1·24	0·57

(continued)

Table 4 (continued)

Variable	n	Unadjusted OR	95% CI	P-value	Adjusted OR	95% CI	P-value
Country of residence							
U.K.	14	–	–		–	–	
Other	19	0·95	0·65–1·37	0·77	0·64	0·32–1·27	0·20
Expert							
No	24	–	–		–	–	
Yes	9	0·87	0·63–1·19	0·38	0·58	0·30–1·14	0·11

OR, odds ratio; CI, confidence interval. [a]Not estimated due to collinearity with site = hands. [b]OR is for every 10% increase in repigmentation.

different from that of the surrounding skin and this will not be captured by the '100% repigmentation' outcome. By focusing on how noticeable the patch is after treatment, the VNS allows factors such as colour match and pattern of repigmentation to be taken into account when assessing treatment success.

It should be noted that the VNS has primarily been developed as a way of assessing treatment response at the end of a period of treatment. It is likely that VNS scores will initially worsen during treatment, as the pigment may return in a darker shade than normal skin (hyperpigmentation). This effect is often temporary, and so it may be preferable to assess VNS scores at a time point after the end of treatment, once residual hyperpigmentation has resolved. The optimal timing for assessing the VNS after treatment will be addressed by further validation work. Nevertheless, having a simple tool that allows quantification of the degree of noticeability during treatment can also help patients to reach informed decisions when choosing between treatment options.

This study is the first to validate an outcome measure assessing treatment success from the patient's perspective. The sample size of > 100 patients provided reasonable precision for estimates of agreement with other measures of treatment success, and specific hypotheses and criteria for success were specified prior to data analysis.[11] We also used the COSMIN checklist to confirm that our study met the COSMIN reporting requirements for validation studies, both for hypothesis testing and interpretability.[15] Our work reports the percentage of missing items, gives a description of how missing items were handled, calculated a sample size, formulated *a priori* hypotheses, gave the expected direction and magnitude of correlations, gave an adequate description of comparator instruments and gave the distribution of total scores. Limitations of the study (with respect to the COSMIN checklist criteria) were that the sample size of clinicians was small, and the comparator instruments (global treatment success and percentage of repigmentation) have not been validated.

While it would be preferable to use actual photographs of vitiligo pre- and post-treatment, limited access to such images necessitated the use of software to simulate repigmentation. Although images attempted to reproduce the patterns of repigmentation typical of those seen after treatment, they were nonetheless simulations. However, this did allow us to control the levels of repigmentation depicted (from fairly minimal to full repigmentation), and ensure complete coverage of the types of repigmentation patterns encountered in a clinical setting. Nonetheless, further validation of the VNS is required, using genuine clinical pre- and post-treatment images.

We focused on 'high-expression' sites (face and hands), as these are areas where vitiligo has the greatest impact.[16] This means that our findings cannot necessarily be applied to vitiligo affecting other anatomical sites. Similarly, the images used in the survey were all of adults with vitiligo, and although some responses were made by the parents/guardians of children with vitiligo, the numbers were so small that we cannot draw firm conclusions about the validity of the outcome measure for assessing treatment response in children. The number of clinician respondents was also relatively small, and may not have been fully representative of all clinicians treating patients with vitiligo. Another limitation of the study is that the participants did not assess their own vitiligo patches. In order to assess the validity of a patient-reported outcome, patients should ideally use the tool on themselves, comparing pretreatment images with the post-treatment clinical appearance; this will be the focus of future work to further validate the scale.

The VNS does not include assessment of physical symptoms such as itch, which can occasionally occur in vitiligo patches. Although these symptoms are not unimportant, the main impact of vitiligo is its visual appearance, so we have not sought to incorporate any physical symptom components into the scale.

Finally, it is possible that the cohort of patient questionnaire respondents is not representative of all individuals who have vitiligo. We used online methods to gather data, and provided the questionnaire only in English. This may have excluded potential respondents from certain ethnic and socioeconomic backgrounds.

This study confirms that the VNS is a valid patient-reported measure of vitiligo treatment success. The VNS is quick and simple to complete, free to use and readily interpreted. However, further validation of the VNS is required, using larger datasets, with genuine clinical pre- and post-treatment images, covering a wider range of anatomical sites and ages of patients, and with patients assessing only their own vitiligo patches, both immediately after treatment and at time intervals after treatment has been completed.

In future clinical trials, patient-reported VNS scores could be used alongside objective outcome measures (such as percentage repigmentation), in order to capture patients' views and to aid interpretation of objective outcome measures. It may also prove to be a useful tool for documentation of treatment response in clinical records.

Acknowledgments

References

1 Ezzedine K, Eleftheriadou V, Whitton M, van Geel N. Vitiligo. *Lancet* 2015; **386**:74–84.

2 Whitton ME, Pinart M, Batchelor J et al. Interventions for vitiligo. *Cochrane Database Syst Rev* 2010; **1**:CD003263.

3 Eleftheriadou V, Whitton ME, Gawkrodger DJ et al. Future research into the treatment of vitiligo: where should our priorities lie? Results of the vitiligo priority setting partnership. *Br J Dermatol* 2011; **164**:530–6.

4 Eleftheriadou V, Thomas KS, Whitton ME et al. Which outcomes should we measure in vitiligo? Results of a systematic review and a survey among patients and clinicians on outcomes in vitiligo trials. *Br J Dermatol* 2012; **167**:804–14.

5 Gonzalez U, Whitton M, Eleftheriadou V et al. Guidelines for designing and reporting clinical trials in vitiligo. *Arch Dermatol* 2011; **147**:1428–36.

6 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**:363–9.

7 Krishna GS, Ramam M, Mehta M et al. Vitiligo impact scale: an instrument to assess the psychosocial burden of vitiligo. *Indian J Dermatol Venereol Leprol* 2013; **79**:205–10.

8 Lilly E, Lu PD, Borovicka JH et al. Development and validation of a vitiligo-specific quality-of-life instrument (VitiQoL). *J Am Acad Dermatol* 2013; **69**:e11–8.

9 Gupta V, Sreenivas V, Mehta M et al. Measurement properties of the Vitiligo Impact Scale-22 (VIS-22), a vitiligo-specific quality-of-life instrument. *Br J Dermatol* 2014; **171**:1084–90.

10 Tour SK, Thomas KS, Walker DM et al. Survey and online discussion groups to develop a patient-rated outcome measure on acceptability of treatment response in vitiligo. *BMC Dermatol* 2014; **14**:10.

11 De Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide.* Cambridge: Cambridge University Press, 2011.

12 Parsad D, Pandhi R, Dogra S, Kumar B. Clinical study of repigmentation patterns with different treatment modalities and their correlation with speed and stability of repigmentation in 352 vitiliginous patches. *J Am Acad Dermatol* 2004; **50**:63–7.

13 Kang C, Qaqish B, Monaco J et al. Kappa statistic for clustered dichotomous responses from physicians and patients. *Stat Med* 2013; **32**:3700–19.

14 Conger AJ. Integration and generalization of kappas for multiple raters. *Psychol Bull* 1980; **88**:322–8.

15 Terwee CB, Mokkink LB, Knol DL et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012; **21**:651–7.

16 Bhandarkar SS, Kundu RV. Quality-of-life issues in vitiligo. *Dermatol Clin* 2012; **30**:255–68.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Summary of image characteristics and distribution of respondents within images.