Further Validation and Responsiveness Assessment of the Herpes Outbreak Impact Questionnaire and Herpes Symptom Checklist

David M. Meads, MSc,¹ Stephen P. McKenna, PhD,¹,² Mary-Ann Bonney, MAppSci,³ Mark T. Bloch, MB, BS, MMed⁴
¹Galen Research, Enterprise House, Manchester Science Park, Lloyd Street North Manchester, UK; ²School of Psychology, University of Central Lancashire, Lancashire, UK; ³Novartis Pharmaceuticals Australia Pty Ltd, North Ryde, NSW, Australia; ⁴Holdsworth House Medical Practice, Darlinghurst, NSW, Australia

ABSTRACT

Objective: The Herpes Outbreak Impact Questionnaire (HOIQ) and Herpes Symptom Checklist (HSC) are patient-completed questionnaires for assessing genital herpes outbreaks. This report describes the validation of paper-and-pencil versions of the scales and responsiveness assessments in an Australian clinical trial.

Methods: Acceptability of the language of the original UK versions was tested with Australians. The HOIQ and HSC were then field-tested with patients. The new versions were validated using patients’ daily responses to the questionnaire in a farniclovir study.

Results: The HOIQ and HSC were readily adapted into Australian English and were acceptable to patients. Psychometric sample: N = 151; 81 (54%) male; mean age 39.9 ± 11.6 years; mean outbreaks 5.1 ± 3.0 per year. Internal consistency was good (alphas at outbreak 1 ranged 0.84–0.90 HOIQ and 0.73–0.87 HSC). Rasch analysis showed item stability over time. Correlations between HOIQ and HSC ranged from 0.46 to 0.60. Both scales distinguished outbreak healing presence or absence at day 6 (P < 0.001), and the HOIQ scale distinguished between symptom severity groups (P < 0.001). Scale scores declined significantly over study duration, exhibiting large effect sizes.

Conclusions: The paper-and-pencil HOIQ and HSC were reliable, valid, and responsive in a clinical trial setting. These instruments are recommended for use in clinical studies.

Keywords: functioning, genital herpes, HOIQ, HSC, patient-reported outcomes, responsiveness, symptoms.

Introduction

High herpes simplex virus prevalence rates have been reported in many developed countries [1], with genital herpes as one of the most common sexually transmitted infections. Although genital herpes is typically asymptomatic for those who experience outbreaks, outbreaks can be painful and debilitating [2]. Recurrent outbreaks are characterized by itching, pain, general flu-like symptoms, and vesicular and ulcerative lesions taking up to 6 days to heal [3]. Additionally, outbreaks can have a significant detrimental psychosocial impact [4].

Psychosocial and quality of life (QoL) impact in recurrent genital herpes (RGH) has been well documented using disease-specific instruments such as the RGHQoL Questionnaire [5]. However, limited information is available on the impact of herpes during an outbreak. The Herpes Outbreak Impact Questionnaire (HOIQ) and the Herpes Symptom Checklist (HSC) were developed to address this need [6].

The HOIQ is a 12-item patient-completed questionnaire assessing the daily impact of herpes outbreaks on the individual’s emotional, self-care, social, sexual, and work-related functioning. Patients rate the impact on a 4-point scale. Total scores range from 0 to 36, in increasing impact on function. The HSC is a 13-item checklist of herpes symptoms, where respondents record the severity of symptoms (tingling, pain, itching, tiredness, etc.) daily on a 4-point scale. Total scores range from 0 to 39, in increasing severity.

Although the scales have been shown to be internally consistent (Cronbach’s alpha: HSC = 0.83; HOIQ = 0.87) and to fit the Rasch model, they were validated via computer administration. The objective of the present study was to provide additional validation of the HOIQ and HSC (via paper-and-pencil administration) and to assess their responsiveness within a clinical trial.

Methods

As the trial was conducted in Australia, it was first necessary to ensure that the UK version of the scales were acceptable to Australians.

Translation

The objective of the “translation” was to produce a conceptually equivalent version for Australians, expressed in everyday language. This was produced using a lay panel of Australians of average or below-average educational level and who did not have a diagnosis of RGH.

Field Test

Field-test interviews were conducted with Australian patients with RGH to assess the relevance, acceptability, comprehensiveness, and clarity of the questionnaire items.

Psychometric Evaluation

Data from a study comparing a 2-day versus 5-day regimen of famciclovir were used to validate the new versions of the measures. Study participants completed the HOIQ and HSC (paper-and-pencil versions) daily for 6 days during two consecutive genital herpes outbreaks. Inclusion criteria for the famciclovir study were: age ≥ 18 years; a clinical diagnosis of genital herpes; and at least two genital herpes outbreaks within the last 12 months, or one outbreak in the last 6 months, or within 6 months...
following an initial episode. Exclusion criteria were: taking suppressive antiviral therapy, pregnancy, breast-feeding, or planning a pregnancy during the study.

Internal consistency was assessed using Cronbach’s alpha. Reproducibility was determined by assessing the stability of the items over time using Rasch analysis [7]. Rasch analysis enables the assessment of differential item functioning (DIF) within the Rasch Unidimensional Measurement Model software framework, with significant DIF taken to indicate item instability.

Convergent validity was assessed by correlating scores on the HOIQ with those on the HSC using Spearman rank correlations. Moderate associations were expected as they measure similar but distinct constructs (impairment and disability).

Discriminative validity was assessed by comparing scores of those whose outbreak had or had not healed according to a physician’s assessment on day 6. Scores on the HOIQ were also examined between individuals who differed according to the level of symptoms they experienced as measured by the HSC. Severity groups were created by using quartile scores on the HSC: 0 to 2, mild symptoms; 3 to 6, moderate symptoms; 7 to 11, severe symptoms; 12+, very severe symptoms.

Differences between the groups were tested by Mann-Whitney U tests or Kruskal–Wallis analyses of variance.

Responsiveness was assessed by effect size thresholds (0.2–0.5 = small change; 0.5–0.8 = moderate change; >0.8 = large change) proposed by Cohen [8] and by the significance of changes in scores from day 1 to day 6 using Wilcoxon signed-rank tests.

Preliminary estimates of the minimal important differences (MIDs) for the scales were derived using distribution-based methods. Scores needed to achieve effect sizes, and the standard error of measurement (SEM) were calculated. Change score needed to achieve a 0.50 effect size (or half a standard deviation) [9], and the SEM [10] have both been suggested as suitable MID estimates.

Results

Translation

The lay panel was comfortable with the HOIQ and HSC wording but changed some words not commonly used in Australia.

Eighteen patients (55.6% male) with RGH, aged 25 to 69 years, participated in field-test interviews. All respondents thought the scales were easily understood and relevant to their experience of genital herpes, although minor changes in wording were adopted. The mean time required to complete the two measures was less than 4 minutes.

Validation

Data from 151 (54% male, mean age 39.9 ± 11.6 years; mean outbreaks per year 5.1 ± 3.0) people were available for the validation and responsiveness analyses (Table 1).

Table 2 shows the range of Cronbach’s alpha coefficients for the HOIQ and HSC (over the 6 days in each outbreak). Both instruments had adequate internal consistency (>0.70); however, the alpha coefficient for the HSC score on day 6 at outbreak 2 was lower—probably because of the low scores, as most outbreaks had healed by this time and reduced sample size on the last day.

No DIF (P > 0.01) by time was found for the HOIQ between day 1 and day 3 scores at outbreak 1 or between day 1 scores of outbreaks 1 and 2. For the same assessment, only one item in the HSC exhibited DIF over time on each occasion. This would not compromise the reproducibility of the HSC. The results provide evidence that the HOIQ and HSC have item stability over time.

A Spearman correlation coefficient matrix was created between HOIQ and HSC scores for each day during outbreak 1. Over the 6 days, the correlations ranged from 0.46 to 0.60, indicating that they are measuring related, but distinct, constructs.

Those who had experienced the condition longer (>sample median time since diagnosis) had fewer symptoms (HSC P = 0.02) and better functioning (HOIQ P = 0.049) at day 6. There were no clear patterns between the number of outbreaks per year and HOIQ or HSC scores.

Individuals whose outbreak had healed by day 6 scored substantially lower on the HOIQ and HSC (P < 0.001) than those whose outbreak had not healed (Fig. 1), providing evidence of the discriminative power of both measures.

The difference in HOIQ scores by symptom group was highly significant (P < 0.001; Fig. 2), confirming that respondents with the most severe symptoms experience the greatest functional limitations.

HOIQ and HSC scores decreased significantly (for both outbreaks) between day 1 and day 6, supporting the responsiveness of the scales (Fig. 3). Effect sizes for the HOIQ and HSC, based on the difference between day 1 and day 6 assessments for both outbreaks, are shown in Table 2. Cohen’s effect size thresholds show that a large change has occurred within both measures in both outbreaks over the complete assessment period, providing further evidence of the responsiveness of these measures.

Distributional estimates of the MIDs (Table 2) suggest a change of between 2 and 2.5 on each scale could be considered important.

Conclusions

Using a lay panel and field tests with relevant patients, new versions of the HOIQ and HSC were created for Australia that were conceptually equivalent to the original UK versions.

Psychometric analyses showed that both measures had adequate internal consistency and that HOIQ and HSC exhibit moderate correlations with each other. Both the HOIQ and HSC exhibit little DIF by time, indicating that responses are not significantly influenced by time and providing evidence of item stability.

Both patient-completed instruments could discriminate outbreak healing in individuals (according to physician assessment)
at their last assessment, providing evidence of construct validity. HOIQ scores differed significantly between symptom severity groups and both scales provided evidence that the impact and severity of herpes outbreaks decrease over time.

A limitation of the study is that, in the absence of suitable RGH outbreak measures, the HOIQ and HSC were correlated with each other. This technique can only provide limited evidence of the validity of either, and in the future, correlational analyses between these and other questionnaires should be conducted. In addition, despite the distributional statistics produced here, further research is required to determine threshold change scores on the questionnaires that could be used to denote an important improvement or a responder.

The psychometric analyses gave promising results in terms of the reliability, validity, and responsiveness of the HOIQ and HSC when administered via paper-and-pencil mode. The validation analyses in this study revealed results similar, in terms of alpha and correlations between the HOIQ and HSC, to those found in the original validation [6]. Although further research with these scales is encouraged, both can be considered suitable for inclusion in clinical trials and for use in clinical practice.

We thank the lay panel, the patients who participated in the field-test interviews, and those in the famciclovir trial. The current research was funded by Novartis Pharmaceuticals Australia Pty Ltd.

Source of financial support: This research was funded by Novartis Pharmaceuticals Australia Pty Ltd.

### Table 2: HOIQ and HSC internal consistency, effect sizes, and distributional minimal important difference

<table>
<thead>
<tr>
<th>Scale</th>
<th>Alpha range (day 1–day 6)</th>
<th>Effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Outbreak 1</td>
<td>Outbreak 2</td>
</tr>
<tr>
<td>HOIQ</td>
<td>0.90–0.84</td>
<td>0.88–0.81</td>
</tr>
<tr>
<td>HSC</td>
<td>0.87–0.73</td>
<td>0.80–0.64</td>
</tr>
</tbody>
</table>

HOIQ, Herpes Outbreak Impact Questionnaire; HSC, Herpes Symptom Checklist.
References