

Published in final edited form as:

Ophthalmology. 2010 December ; 117(12): 2322–2328.e1. doi:10.1016/j.ophtha.2010.03.042.

Responsiveness of health-related quality of life questionnaires in adults undergoing strabismus surgery

Sarah R. Hatt, D.B.O.¹, David A. Leske, M.S.¹, and Jonathan M. Holmes, B.M., B.Ch.¹

¹Department of Ophthalmology, Mayo Clinic, Rochester, MN

Abstract

Objective—To evaluate the responsiveness of 2 health related quality of life (HRQOL) questionnaires, the Adult Strabismus-20 (AS-20) and National Eye Institute Visual Function Questionnaire-25 (VFQ-25), to changes in HRQOL following strabismus surgery.

Design—Cohort study

Participants—106 adults (aged 18 to 84, median 48.5 years) undergoing strabismus surgery, 80 (75%) with diplopia and 26 (25%) without diplopia.

Methods—All participants completed AS-20 and VFQ-25 questionnaires pre-operatively and a median of 7 (range 4 to 13) weeks post-operatively. Using pre-defined clinical criteria, post-operative outcomes were graded as either ‘success’ (n=65), ‘partial success’ (n=32) or ‘failure’ (n=9).

Main outcome measures—For each questionnaire, mean composite and subscale scores (ranging from 0 to 100; worst to best HRQOL) were calculated for each patient. Median change in scores pre- to post-operatively was compared for successes, partial successes, and failures, and the proportion of patients whose scores exceeded 95% limits of agreement was calculated.

Results—Successfully aligned diplopic patients (n=46) showed greater improvement than failures (n=7) using both the AS-20 questionnaire (21.3 [quartiles 12.5 to 35.0] vs. 8.8 [3.8 to 12.5]; P=0.002) and the VFQ-25 questionnaire (18.3 [8.6 to 26.1] vs. 8.3 [0.8 to 13.5]; P=0.02). Successfully aligned non-diplopic patients (n=19) also showed greater improvement than failures (n=2): (AS-20 23.8 [10.0 to 32.5] vs. -3.1 [-10.0 to 3.8]; P=0.05). In non-diplopic patients, changes on VFQ-25 were small but significantly greater for successes than failures (5.0 [0.0 to 10.1] vs. -15.4 [-19.4 to -11.4]; P=0.03). More successfully aligned patients showed improvement exceeding the 95% limits of agreement for AS-20 scores than VFQ-25 scores (67% vs. 65% [P=0.8] for diplopic patients and 63% vs. 21% [P=0.005] for non-diplopic patients).

Conclusions—AS-20 and VFQ-25 questionnaires are responsive to improved HRQOL in adults undergoing successful strabismus surgery. Changes on VFQ-25 were smaller, particularly for non-

© 2010 American Academy of Ophthalmology, Inc. Published by Elsevier Inc. All rights reserved.

Correspondence and reprint requests to: Dr. Jonathan M. Holmes, Ophthalmology W7, Mayo Clinic, Rochester, MN 55905. Phone: (507) 284-3760. Fax: (507) 284-8566. holmes.jonathan@mayo.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Presented in part at: The Association for Research in Vision and Ophthalmology meeting, Fort Lauderdale, FL, USA, May 4th 2009, and at the European Strabismological Association meeting, Belgrade, Serbia, October 7th 2009.

This article contains additional online-only material. The following should appear online-only: Table 4.

diplopic strabismus. The AS-20 is more responsive than the VFQ-25 across the range of adult strabismus.

Treatment for adult strabismus aims to address effects on an individual's health related quality of life (HRQOL), but formal evaluation of HRQOL is rarely performed. Several HRQOL instruments have been used to evaluate the effects of strabismus in adults.¹⁻⁷ These previous studies have reported concerns such as low self-esteem, problems with interpersonal relationships and social anxiety in adults with strabismus. Nevertheless few studies report prospective evaluation of HRQOL before and after treatment. The ability of a HRQOL instrument to reflect clinically significant change either over time, or following treatment, is termed 'responsiveness' and is an important aspect of instrument validity.⁸⁻¹⁰ In previous reports we have described the development, initial validation, and test retest reliability of the Adult Strabismus-20 (AS-20) questionnaire, a strabismus-specific HRQOL questionnaire for adults.^{7,11-13} The AS-20 was developed to measure the specific concerns of strabismus patients, which we suspected were not fully addressed in existing vision-related HRQOL questionnaires, such as the widely used National Eye Institute Visual Function Questionnaire-25 (VFQ-25).¹⁴ When we previously compared the AS-20 and VFQ-25 in a cohort of adult strabismus patients, we found the AS-20 was more sensitive in detecting reduced HRQOL than the VFQ-25, particularly in patients with non-diplopic strabismus.¹² In this present study, we further evaluated the validity of the AS-20 questionnaire by analyzing responsiveness to changes in ocular alignment and symptoms following strabismus surgery. We also assessed the responsiveness of the VFQ-25¹⁴ in the same patients.

Patients and Methods

Approval was obtained from the Institutional Review Board, Mayo Clinic, Rochester MN and each patient gave informed consent before participating. All procedures and data collection were conducted in a manner compliant with the Health Insurance Portability and Accountability Act.

Patients

One hundred and twelve consecutive adult strabismus patients were prospectively recruited from the outpatient clinics of one author's practice (JM) and completed AS-20 and VFQ-25 questionnaires pre- and post-operatively. Patients were not recruited if they could not read or understand English or if they had severe cognitive impairment. Patients with ocular myasthenia gravis were excluded because the condition may vary from day to day. For each patient, angle of deviation was measured pre- and post-operatively, at distance and near, using simultaneous prism and cover test (SPCT) and prism and alternate cover test (PACT). Eighty (71%) of 112 had diplopia, 26 (23%) had no diplopia, and 6 (5%) had rare diplopia (occasional diplopia by history; no diplopia on exam). We excluded patients with rare diplopia from further analysis because there were only 6 and these patients may behave differently from either the diplopic or non-diplopic patients. For the remaining 106 patients, median age was 48.5 (range 18 to 84) years. Sixty (57%) were female and 98 (92%) self-reported their race as 'White.' Forty-nine (46%) had childhood onset / idiopathic strabismus, 33 (31%) neurogenic, 14 (13%) mechanical, and 10 (9%) sensory. Pre-operatively patients were classified by the direction of their deviation. For multi-planar deviations, classification was by the largest of the horizontal or vertical deviation. Thirty-nine (37%) primarily had an esodeviation, 43 (41%) an exodeviation, and 24 (23%) a vertical deviation. For 5 of 26 non-diplopic patients, Krinsky measurements were used as visual acuity was too poor to allow accurate cover testing. Median visual acuity was 20/20 (range 20/15 to 20/40) in the better eye and 20/25 (range 20/15 to hand motions) in the worse eye. Twenty-three (22%) had ocular co-morbidity such as glaucoma or corneal abnormalities. Forty-six (43%) had

undergone strabismus surgery previously in a separate episode of care. Assessments were made a median of one day (range 1 to 27 days) pre-operatively and 7 weeks (range 4 to 13 weeks) post-operatively. Pre-operative data have been reported on 44 (42%) of the 106 patients in previous analyses^{7, 13} and for 40 of these, data were from the same pre-operative examination as the current report. None of the post-operative data have been previously reported.

Questionnaire administration

Patients completed AS-20 and VFQ-25 questionnaires at pre- and post-operative assessments. Questionnaires were self-administered, without supervision, following simple verbal and written instructions. Patients were instructed to respond as when wearing any habitual refractive correction. To standardize testing procedures, AS-20 questionnaires were completed before VFQ-25 questionnaires.

AS-20 questionnaire

The AS-20 consists of a total of 20 items, 10 in a psychosocial subscale and 10 in a function subscale (full questionnaire with user instructions available at: <http://public.pedig.jaeb.org/>; accessed January 27, 2010). For each question, a 5-point Likert type scale is used for responses: 'never' (score 100), 'rarely' (score 75), 'sometimes' (score 50), 'often' (score 25), and 'always' (score 0).⁷ AS-20 composite score (20 items), psychosocial subscale score (10 items), and function subscale score (10 items) range from 0 to 100 (worst to best HRQOL) and are calculated as a mean of all answered items.⁷ Of 106 patients, one diplopic patient failed to complete the AS-20 psychosocial subscale, preventing calculation of a psychosocial subscale score.

VFQ-25 questionnaire

The VFQ-25 consists of 25 items in 12 subscales: general health, general vision, ocular pain, near activities, distance activities, vision-specific social function, vision-specific mental health, vision-specific role difficulties, vision-specific dependency, driving, color vision, and peripheral vision (full questionnaire with user instructions available at: http://www.nei.nih.gov/resources/visionfunction/vfq_ia.pdf; accessed January 27, 2010). VFQ-25 composite and subscale scores range from 0 to 100 (worst to best HRQOL) and are calculated as a mean of all answered items. As instructed in the manuscript first describing the VFQ-25,¹⁴ the composite score is calculated as a mean of all subscales, with the exception of the 'General Health' subscale. Due to incomplete data, the driving subscale score could not be calculated in 5 patients (3 diplopic and 2 non-diplopic), the color vision subscale in 2 diplopic patients, the vision-specific role difficulties subscale in 1 diplopic patient, the vision-specific dependency subscale in 1 diplopic patient, and the near activities subscale in 1 non-diplopic patient..

Classification of post-operative outcomes

For this study of responsiveness, we used clinical criteria, defined a priori, to classify post-operative outcomes as either 'success,' 'partial success,' or 'failure.' For 'success,' all of the following four criteria had to be met: 1) no diplopia / visual confusion in primary position (distance or near) and for reading (by history and by exam); 2) <10 pd heterotropia in primary position at distance and near by SPCT; 3) not using prism / Bangerter foil / occlusion; 4) no other symptoms potentially related to misalignment or strabismus surgery, e.g. eyestrain, blurry vision, photophobia, or suture reaction. Patients categorized 'partial success' met all of the following criteria: 1) no diplopia / visual confusion in primary position or no diplopia / visual confusion for reading; 2) <20 pd heterotropia in primary position at distance and near by SPCT; 3) not using prism / Bangerter foil / occlusion; 4) any

symptoms such as eyestrain, blurry vision, photophobia, or suture reaction present only mildly or intermittently. Patients were classified as surgical 'failure' if any one of the following criteria were met: 1) diplopia / visual confusion in primary position and reading; 2) >20 pd heterotropia in primary position at distance or near by SPCT; 3) using prism / Bangerter foil / occlusion; 4) moderate to severe symptoms related to misalignment or strabismus surgery, such as eyestrain or blurry vision.

Analyses

For each questionnaire, mean composite and subscale scores were calculated for each patient. Pre- to post-operative change in composite and subscale scores was analyzed in three ways. The analysis of primary interest was the comparison of median change in scores in successes versus failures. 1) Median change in scores was compared between outcome groups (success, partial success, and failure) for patients with and without diplopia, using Wilcoxon tests. Non-parametric tests were used because data were not normally distributed. 2) Using previously published¹³ 95% limits of agreement (LOA) for AS-20 composite score (14.3), VFQ-25 composite score (11.1), AS-20 subscales and VFQ-25 subscales, the proportions of patients showing change exceeding the 95% LOA were compared for AS-20 and VFQ-25 composite scores using McNemar's test. Proportions were also compared across groups (success, partial success, and failure) using Fishers exact test for patients with and without diplopia. 3) The effect size statistic was calculated by dividing mean change in scores (post-operative score minus preoperative score) by the standard deviation (SD) of baseline scores.¹⁰ A higher effect size indicates a more responsive measure; effect sizes of 0.20 to 0.49 are considered small, 0.50 to 0.79 are considered medium, and 0.80 and higher are considered large.¹⁵

Pre-operative composite and subscale scores were compared between diplopic and nondiplopic patients and between those with and without previous surgery, using Wilcoxon rank tests.

All statistical analyses were performed using SAS computer software version 9.1.3.

Results

Applying our pre-specified outcome criteria, 65 (61%) patients were classified 'success' (46 with diplopia, 19 without diplopia), 32 (30%) 'partial success' (27 with diplopia, 5 without diplopia), and 9 (8%) 'failure' (7 with diplopia, 2 without diplopia).

Pre-operative questionnaire scores

Comparing pre-operative composite scores for patients with and without diplopia, median (1st quartile, 3rd quartile) AS-20 scores were 57.5 (45.6, 70.6) for diplopic patients and 61.3 (42.5, 70.0) for non-diplopic patients ($P=1.0$). For the VFQ-25, median pre-operative scores were 71.5 (55.6, 82.0) for diplopic patients and 88.3 (83.1, 91.6) for non-diplopic patients ($P<0.0001$).

Comparing AS-20 subscales between diplopia and non-diplopic patients, median psychosocial scores were lower in non-diplopic patients (50.0 vs. 72.5; $P=0.001$) whereas median function scores were lower in diplopic patients (42.5 vs. 66.3; $P=0.0002$).

Comparing VFQ-25 subscales between diplopic and non-diplopic patients, the following were significantly different with lower median scores for diplopic patients: general vision (60.0 vs 80.0; $P=0.009$), near activities (66.7 vs 91.7; $P<0.0001$), distance activities (66.7 vs 91.7; $P<0.0001$), vision specific social function (87.5 vs 100.0; $P=0.03$), vision specific mental health (56.3 vs 81.3; $P=0.0003$), vision specific role difficulties (50.0 vs 87.5; $P<0.0001$), vision specific dependency (83.3 vs 100.0; $P=0.0005$), driving (75.0 vs 83.3;

P=0.009) and peripheral vision (62.5 vs 100.0; P=0.0007). There were no differences between diplopic and non-diplopic patients on the general health, ocular pain and color vision subscales (P>0.05 for each comparison).

Comparing pre-operative scores between patients who had undergone previous surgery (n=46) and those who had not (n=60), scores were similar (AS-20: 61.3 vs. 56.9; P=0.9; VFQ-25: 81.9 vs. 73.2; P=0.052). All subscale scores were also similar between those with and without previous surgery, with the exception of the VFQ-25 vision-specific dependency subscale which was higher in those who had undergone previous surgery (100.0 vs. 83.3; P=0.04).

Average change in scores from pre- to post-operatively

Patients with diplopia—Median change in AS-20 composite score was significantly greater for successes (21.3 [quartiles 12.5, 35.0]) compared with failures (8.8 [3.8, 12.5]; P=0.002) and also compared with partial successes (17.5 [6.3, 22.5]; P=0.05; Figure 1A, Table 1). For AS-20 subscales, change in the function subscale was significantly different between successes and failures (27.5 [20.0, 47.5] vs. 7.5 [5.0, 17.5]; P=0.004), successes and partial successes (20 [12.5, 35.0]; P=0.01), and partial successes and failures (P=0.04; Table 1). For the AS-20 psychosocial subscale, differences between successes and failures (10.0 [2.5, 27.5] vs. 0.0 [-2.5, 17.5]) approached significance (P=0.052; Table 1).

For the VFQ-25 composite score, median change was significantly greater for successes (18.3 [8.6, 26.1]) compared with failures (8.3 [0.8, 13.5]; P=0.02) and also compared with partial successes (10.7 [4.0, 18.7], P=0.02; Figure 1B, Table 1). Comparisons of VFQ-25 subscales are shown in Table 1.

Patients without diplopia—Median change in AS-20 composite score was significantly greater for successes (23.8 [10.0, 32.5]) compared with failures (-3.1 [-10.0, 3.8]; P=0.05; Figure 2A, Table 2). Changes on the AS-20 psychosocial subscale and the function subscale were not significantly different between successes, partial successes, and failures (P>0.07 for each comparison; Table 2).

For the VFQ-25 composite score, median change in scores was significantly greater for successes (5.0 [0.0, 10.1]) compared with failures (-15.4 [-19.4, -11.4]; P=0.03; Figure 2B, Table 2). Comparisons of VFQ-25 subscales are shown in Table 2.

Limits of agreement

Patients with diplopia—Comparing questionnaires, similar proportions of diplopic patients showed improvement exceeding the 95% LOA on AS-20 and VFQ-25 composite scores in successes (67% vs. 65%, P=0.8) and in partial successes (56% vs. 48%, P=0.5; Table 3). Comparing successes and failures, a greater proportion of successfully aligned patients showed improvement in AS-20 scores exceeding 95% LOA when compared with failures (67% vs. 0%, P=0.001; Table 3) whereas similar proportions of successes and failures showed improvement in VFQ-25 (65% vs. 43%; P=0.4). On both AS-20 and VFQ-25, one patient (partial success) showed a decrease in score (worsening) exceeding 95% LOA.

Patients without diplopia—Comparing questionnaires, more of the successfully aligned non-diplopic patients showed improvement exceeding the 95% LOA on AS-20 composite score than on VFQ-25 composite score (63% vs. 21%, P=0.005; Table 3). Comparing successes and failures, a numerically greater proportion of successfully aligned patients showed improvement in AS-20 scores exceeding the LOA when compared with failures

(63% vs. 0% $P=0.2$; Table 3) whereas similar proportions of successes and failures showed improvement in VFQ-25 (21% vs. 0%; $P=1.0$). No patients showed a decrease (worsening) in score exceeding 95% LOA on AS-20 composite score, but the two failure patients showed worsening on VFQ-25 composite score.

Effect size

Patients with diplopia—For the AS-20 composite score effect size was 1.27 in successes, 0.88 in partial successes, and 0.47 in failures (Table 4; available at <http://aaojournal.org>). For successfully aligned patients, the effect size was greater on the AS-20 function subscale than on the psychosocial subscale (1.56 vs. 0.68) (Table 4; available at <http://aaojournal.org>). For the VFQ-25 composite score, the effect size was 1.14 in successes, 0.69 in partial successes, and 0.54 in failures (Table 4; available at <http://aaojournal.org>). In successfully aligned patients, the VFQ-25 subscale with the highest effect size was the vision specific mental health subscale (1.18).

Patients without diplopia—For the AS-20 composite score, the effect size was 1.43 in successes, 0.99 in partial successes, and -0.18 in failures (Table 4; available at <http://aaojournal.org>). For successfully aligned patients, effect size was greater on the AS-20 psychosocial subscale than on the function subscale (1.44 vs. 0.70) (Table 4; available at <http://aaojournal.org>). In successes, the effect size was low for VFQ-25 composite scores (0.57, Table 4; available at <http://aaojournal.org>). None of the VFQ-25 subscales showed large effect sizes (highest = 0.76 for the vision specific mental health subscale).

Discussion

Both AS-20 and VFQ-25 questionnaires were responsive to change in ocular alignment and symptoms following strabismus surgery, showing significantly greater post-operative improvement in successfully aligned patients compared with partial successes and surgical failures. Across the three methods used to assess responsiveness, the AS-20 questionnaire was more responsive than the VFQ-25 for non-diplopic patients.

Although responsiveness has been reported for HRQOL instruments used in other eye diseases,¹⁶ there are few data on instruments used in adult strabismus. Jackson et al⁵ used three well established, generic measures of social anxiety, depression, and quality of life (Hospital Anxiety and Depression Scale [HADS], Derriford Appearance Scale and WHOQoLBref), and found that all three were responsive in adults undergoing strabismus surgery. Nevertheless, in that study, change in questionnaire scores was not interpreted in the context of specific categories of success or failure, and it is possible that apparent responsiveness was due to a placebo effect from surgery. By carefully defining criteria for post-surgical success, partial success, and failure, we aimed to reduce the influence of a possible placebo effect on our findings. In a recent study by Chai et al,¹⁷ the VFQ-25 and HADS questionnaires (Chinese versions) were administered as proxy measures to parents of children undergoing strabismus surgery. Improvements in HRQOL were reported using both questionnaires, but the VFQ-25 was designed for adults and not intended for use as a proxy questionnaire. In other studies, Beauchamp et al³ reported pre- and post-operative HRQOL using a disability questionnaire, but pre-operative questionnaires were completed retrospectively, at least 6 months after surgery. In our present study, we administered AS-20 and VFQ-25 questionnaires prospectively and defined outcomes using pre-specified clinical measures.

We previously reported that the AS-20 was more sensitive than the VFQ-25 in detecting reduced HRQOL in adult strabismus, especially in patients without diplopia.¹² Although the analysis of median change in VFQ-25 in the present study showed some responsiveness in

non-diplopic patients, many VFQ-25 scores in such patients were not subnormal pre-operatively (based on previous thresholds¹²) and therefore had little room for improvement. The difference between non-diplopic and diplopic patients in pre-operative VFQ-25 scores appears to be driven by relatively normal VFQ-25 scores in non-diplopic patients. Ideally, a HRQOL instrument would indicate poorer scores in an affected population than in controls, but non-diplopic patients often have normal scores on the VFQ-25.¹²

Despite some responsiveness of the VFQ-25 based on median change in score, overall there was poor responsiveness based on limits of agreement and effect size. Nevertheless, some VFQ-25 subscales showed good responsiveness across different methods of analysis, in particular, the vision specific mental health subscale, comprising four questions, addressing vision-related worry, frustration, self-consciousness and 'having less control over what I do.' Interestingly, in the initial development of the AS-20, we identified these same concerns,¹¹ many of which are represented in the final AS-20 questionnaire. The VFQ-25 vision specific mental health subscale may compliment the AS-20 questionnaire by addressing concerns pertinent to adults with strabismus.

By categorizing patients post-operatively as success, partial success, or failure according to pre-specified clinical criteria, data in the present study provide evidence of construct validity for each questionnaire, demonstrating that questionnaire scores change in an expected way when the condition improves or deteriorates. We tested the theoretical construct that HRQOL scores would change more in patients judged to be successfully aligned than in patients who were unsuccessfully aligned. Using three different methods of analysis, the AS-20 questionnaire responded as expected in diplopic and non-diplopic patients whereas the VFQ-25 did not respond as well in non-diplopic patients. Interestingly, not only did the surgical failures improve less than surgical successes but, in patients without diplopia, many scores actually decreased post-operatively. Nevertheless, this observation is limited by the fact that there were only two non-diplopic patients who were classified as failures.

Our finding of improved HRQOL in adults following successful strabismus surgery has implications for clinical management. As reported in previous studies, strabismus often has significant effects not only on physical functioning but also on psychosocial functioning,^{1-5, 11, 18} and formal assessment of HRQOL is one method of defining and quantifying the nature and severity of a patient's concerns. The AS-20 appears to be particularly useful because it evaluates, and is responsive to, both physical and psychosocial aspects of strabismus.

There are some potential weaknesses to this study. We had fewer non-diplopic patients and few failures among diplopic and non-diplopic patients, which may have limited our ability to find differences between successes, partial successes, and failures in some analyses. Also, to reflect common clinical practice, we used relatively early post-operative examinations to assess outcomes (median 7 weeks). It is possible that effects on HRQOL continue to change over a more extended period of follow up, and we aim to evaluate this in future studies. Regarding methods used to assess responsiveness, it is possible that an average improvement in HRQOL scores may occur due to regression to the population mean, a function of simply repeating the measurement, but we feel that this is unlikely since there was significantly less improvement in patients for whom surgery was unsuccessful. In addition, we have previously reported negligible regression to the mean for repeat administration of AS-20 and VFQ-25 over time.¹³

Both AS-20 and VFQ-25 HRQOL questionnaires were responsive to changes in ocular alignment in adults undergoing successful strabismus surgery, although the condition-specific AS-20 questionnaire showed greater responsiveness than the VFQ-25 in non-

diplopic strabismus. The AS-20 questionnaire is a useful tool for assessing the impact of strabismus on HRQOL in adults and for measuring changes in HRQOL following surgery.

Acknowledgments

Supported by National Institutes of Health Grant EY015799 (JMH), Research to Prevent Blindness, New York, NY (JMH as Olga Keith Weiss Scholar and an unrestricted grant to the Department of Ophthalmology, Mayo Clinic), and Mayo Foundation, Rochester, MN.

None of the funding organizations had any role in the design or conduct of this research.

References

1. Satterfield D, Keltner JL, Morrison TL. Psychosocial aspects of strabismus study. *Arch Ophthalmol* 1993;111:1100–5. [PubMed: 8166786]
2. Menon V, Saha J, Tandon R, et al. Study of the psychosocial aspects of strabismus. *J Pediatr Ophthalmol Strabismus* 2002;39:203–8. [PubMed: 12148552]
3. Beauchamp GR, Black BC, Coats DK, et al. The management of strabismus in adults--III. The effects on disability. *J AAPOS* 2005;9:455–9. [PubMed: 16213395]
4. Burke JP, Leach CM, Davis H. Psychosocial implications of strabismus surgery in adults. *J Pediatr Ophthalmol Strabismus* 1997;34:159–64. [PubMed: 9168420]
5. Jackson S, Harrad RA, Morris M, Rumsey N. The psychosocial benefits of corrective surgery for adults with strabismus. *Br J Ophthalmol* 2006;90:883–8. [PubMed: 16782950]
6. van de Graaf ES, van der Sterre GW, Polling JR, et al. Amblyopia & Strabismus Questionnaire: design and initial validation. *Strabismus* 2004;12:181–93. [PubMed: 15370526]
7. Hatt SR, Leske DA, Bradley EA, et al. Development of a quality-of-life questionnaire for adults with strabismus. *Ophthalmology* 2009;116:139–44. [PubMed: 19019449]
8. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171–8. [PubMed: 3818871]
9. Guyatt GH, Deyo RA, Charlson M, et al. Responsiveness and validity in health status measurement: a clarification. *J Clin Epidemiol* 1989;42:403–8. [PubMed: 2659745]
10. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures: statistics and strategies for evaluation. *Control Clin Trials* 1991;12(suppl):142S–58S. [PubMed: 1663851]
11. Hatt SR, Leske DA, Kirgis PA, et al. The effects of strabismus on quality of life in adults. *Am J Ophthalmol* 2007;144:643–7. [PubMed: 17707329]
12. Hatt SR, Leske DA, Bradley EA, et al. Comparison of quality-of-life instruments in adults with strabismus. *Am J Ophthalmol* 2009;148:558–62. [PubMed: 19570519]
13. Leske DA, Hatt SR, Holmes JM. Test-retest reliability of health-related quality-of-life questionnaires in adults with strabismus. *Am J Ophthalmol*. In press.
14. Mangione CM, Lee PP, Gutierrez PR, et al. National Eye Institute Visual Function Questionnaire Field Test Investigators. Development of the 25-item National Eye Institute Visual Function Questionnaire. *Arch Ophthalmol* 2001;119:1050–8. [PubMed: 11448327]
15. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed.. L. Erlbaum Assoc.; Hillsdale, NJ: 1988. p. 19-74.
16. Margolis MK, Coyne K, Kennedy-Martin T, et al. Vision-specific instruments for the assessment of health-related quality of life and visual functioning: a literature review. *Pharmacoeconomics* 2002;20:791–812. [PubMed: 12236802]
17. Chai Y, Shao Y, Lin S, et al. Vision-related quality of life and emotional impact in children with strabismus: a prospective study. *J Int Med Res* 2009;37:1108–14. [PubMed: 19761693]
18. Nelson BA, Gunton KB, Lasker JN, et al. The psychosocial aspects of strabismus in teenagers and adults and the impact of surgical correction. *J AAPOS* 2008;12:72–6. [PubMed: 18314071]

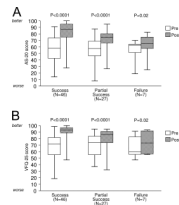


Figure 1.

Patients with diplopia. Median Adult Strabismus-20 (AS-20) composite scores (A) and National Eye Institute Visual Function Questionnaire-25 (VFQ-25) composite scores (B) at pre-operative and post-operative exams by post-operative status (successfully aligned, partial success, and failure). Boxes represent 1st quartile, median and 3rd quartile values; whiskers represent extreme values. All comparisons between pre- and post-surgery scores were significant for AS-20 and VFQ-25 (P values shown). For the AS-20 (A), change was greater for successes than failures and partial successes. For the VFQ-25 (B), change was greater for successes than failures and partial successes.

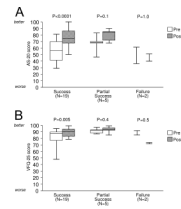


Figure 2. Patients without diplopia. Median Adult Strabismus-20 (AS-20) composite scores (A) and National Eye Institute Visual Function Questionnaire-25 (VFQ-25) composite scores (B) at pre-operative and post-operative exams by post-operative status (successfully aligned, partial success, and failure). Boxes represent 1st quartile, median, and 3rd quartile values; whiskers represent extreme values. The only pre- to postoperative comparisons that were significant were among successes (P values shown). For the AS-20 (A), change was greater for successes than failures. For the VFQ-25 (B), change was greater for successes than failures.

Table 1
Health related quality of life scores in patients with diplopia (n=80) using AS-20 and VFQ-25 pre and post-operatively.

	Diplopic patients (n=80)					
	Pre-Post op difference SUCCESS (n=46)	Pre-Post op difference PARTIAL SUCCESS (n=27)	Pre-Post op difference FAILURE (n=7)	Success vs Failure	Success vs Partial success	Partial success vs failure
	Median (1 st quartile, 3 rd quartile)	Median (1 st quartile, 3 rd quartile)	Median (1 st quartile, 3 rd quartile)			
AS-20 questionnaire						
AS-20 composite score	21.3 (12.5, 35.0)	17.5 (6.3, 22.5)	8.8 (3.8, 12.5)	0.002	0.05	0.06
Psychosocial subscale	10.0 (2.5, 27.5)	7.5 (2.5, 20.0)	0.0 (-2.5, 17.5)	0.052	0.3	0.2
Function subscale	27.5 (20.0, 47.5)	20.0 (12.5, 35.0)	7.5 (5.0, 17.5)	0.004	0.01	0.04
VFQ-25 questionnaire						
VFQ-25 composite score	18.3 (8.6, 26.1)	10.7 (4.0, 18.7)	8.3 (0.8, 13.5)	0.02	0.02	0.7
General health	0.0 (0.0, 0.0)	0.0 (-25.0, 25.0)	0.0 (-25.0, 25.0)	0.8	0.4	0.9
General vision	20.0 (0.0, 40.0)	20.0 (0.0, 20.0)	0.0 (0.0, 20.0)	0.2	0.054	0.8
Ocular pain	12.5 (0.0, 25.0)	0.0 (0.0, 12.5)	0.0 (0.0, 12.5)	0.07	0.06	0.7
Near activities	16.7 (8.3, 41.7)	8.3 (0.0, 25.0)	8.3 (-8.3, 25.0)	0.2	0.1	0.7
Distance activities	16.7 (8.3, 33.3)	8.3 (0.0, 25.0)	0.0 (-8.3, 16.7)	0.03	0.1	0.2
Vision specific social functioning	12.5 (0.0, 25.0)	0.0 (0.0, 25.0)	12.5 (0.0, 12.5)	0.9	0.1	0.5
Vision specific mental health	31.3 (12.5, 50.0)	18.8 (6.3, 37.5)	6.3 (-12.5, 12.5)	0.001	0.05	0.02
Vision specific role difficulties	25.0 (0.0, 50.0)	12.5 (0.0, 25.0)	12.5 (0.0, 25.0)	0.3	0.08	1.0
Vision specific dependency	12.5 (0.0, 25.0)	8.3 (0.0, 25.0)	0.0 (0.0, 8.3)	0.2	0.5	0.3
Driving	16.7 (0.0, 25.0)	8.3 (0.0, 16.7)	4.2 (0.0, 25.0)	0.3	0.1	0.9
Color vision	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 25.0)	0.1	0.2	0.04
Peripheral vision	25.0 (0.0, 50.0)	0.0 (0.0, 25.0)	25.0 (0.0, 50.0)	1.0	0.1	0.5

Changes in scores for successes, partial successes and failures represented as medians (quartiles). Differences between successes, partial successes and failures were analyzed using Wilcoxon tests. AS-20 = Adult Strabismus-20; VFQ-25 = National Eye Institute Visual Function Questionnaire-25

Table 2
Health related quality of life scores in patients without diplopia (n=26) using AS-20 and VFQ-25 pre and post-operatively.

	Non-diplopic patients (n=26)					
	Pre-Post op difference SUCCESS (n=19)	Pre-Post op difference PARTIAL SUCCESS (n=5)	Pre-Post op difference FAILURE (n=2)	Success vs Failure	Success vs Partial success	Partial success vs failure
	Median (1 st quartile, 3 rd quartile)	Median (1 st quartile, 3 rd quartile)	Median (1 st quartile, 3 rd quartile)			
AS-20 questionnaire						
AS-20 composite score	23.8 (10.0, 32.5)	2.5 (2.5, 22.0)	-3.1 (-10, 3.8)	0.05	0.3	0.6
Psychosocial subscale	27.5 (15.0, 47.5)	5.0 (2.5, 14.7)	2.5 (-2.5, 7.5)	0.07	0.1	0.8
Function subscale	7.5 (2.5, 25.0)	10.0 (0.0, 12.5)	-8.8 (-27.5, 10.0)	0.3	0.6	0.4
VFQ-25 questionnaire						
VFQ-25 composite score	5.0 (0.0, 10.1)	1.8 (0.9, 2.9)	-15.4 (-19.4, -11.4)	0.03	0.3	0.08
General health	0.0 (0.0, 0.0)	0.0 (0.0, 25.0)	25.0 (0.0, 50.0)	0.3	0.9	0.5
General vision	0.0 (0.0, 20.0)	20.0 (0.0, 20.0)	0.0 (-20.0, 20.0)	0.7	0.5	0.7
Ocular pain	0.0 (-12.5, 25.0)	0.0 (0.0, 0.0)	-18.8 (-37.5, 0.0)	0.2	0.7	0.4
Near activities	0.0 (0.0, 8.3)	0.0 (-4.2, 4.2)	-16.7 (-50, 16.7)	0.8	0.3	1.0
Distance activities	0.0 (0.0, 25.0)	0.0 (-8.3, 0.0)	-4.2 (-8.3, 0.0)	0.2	0.06	1.0
Vision specific social functioning	0.0 (0.0, 25.0)	0.0 (0.0, 0.0)	-25.0, (-25.0, -25.0)	0.02	0.1	0.03
Vision specific mental health	12.5 (0.0, 18.8)	0.0 (-6.3, 0.0)	-17.7 (-18.8, -16.7)	0.03	0.03	0.2
Vision specific role difficulties	12.5 (0.0, 12.5)	0.0 (0.0, 0.0)	-37.5 (-37.5, -37.5)	0.04	0.6	0.052
Vision specific dependency	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	-16.7 (-25.0, -8.3)	0.007	0.8	0.055
Driving	0.0 (0.0, 16.7)	0.0 (0.0, 0.0)	-8.3 (-8.3, -8.3)	0.05	0.3	0.3
Color vision	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	-12.5 (-25.0, 0.0)	0.03	0.7	0.2
Peripheral vision	0.0 (0.0, 25.0)	0.0 (0.0, 25.0)	-12.5 (-25.0, 0.0)	0.3	0.4	0.2

Changes in scores for successes, partial successes and failures represented as medians (quartiles). Differences between successes, partial successes and failures were analyzed using Wilcoxon tests. AS-20 = Adult Strabismus-20; VFQ-25 = National Eye Institute Visual Function Questionnaire-25

Table 3
 Proportion of adult strabismus patients showing pre- to post-operative change in health related quality of life scores exceeding previously defined limits of agreement.

AS-20 questionnaire	Patients showing change exceeding 95% limits of agreement: n (%)					
	Diplopia (n=80)			No diplopia (n=26)		
	Success (n=46)	Partial Success (n=27)	Failure (n=7)	Success (n=19)	Partial Success (n=5)	Failure (n=2)
AS-20 composite score	31 (67%) ^a	15 (56%) ^c	0 (0%) ^{a, c}	12 (63%) ^d	2 (40%)	0 (0%)
Psychosocial subscale	18 (39%)	7 (26%)	1 (14%)	13 (68%)	1 (20%)	0 (0%)
Function subscale	35 (76%) ^{a, b}	14 (52%) ^b	1 (14%) ^d	5 (26%)	1 (20%)	0 (0%)
VFQ-25 questionnaire						
VFQ-25 composite score	30 (65%)	13 (48%)	3 (43%)	4 (21%) ^d	0 (0%)	0 (0%)
General health	8 (17%)	7 (26%)	2 (29%)	4 (21%)	2 (40%)	1 (50%)
General vision	15 (33%)	4 (15%)	1 (14%)	2 (11%)	0 (0%)	0 (0%)
Ocular pain	9 (20%)	3 (11%)	0 (0%)	1 (5%)	0 (0%)	0 (0%)
Near activities	22 (48%)	8 (30%)	2 (29%)	2 (11%)	0 (0%)	0 (0%)
Distance activities	22 (48%)	7 (26%)	1 (14%)	5 (26%)	0 (0%)	0 (0%)
Vision specific social functioning	16 (35%)	8 (30%)	1 (14%)	5 (26%)	0 (0%)	0 (0%)
Vision specific mental health	26 (57%) ^a	10 (37%)	0 (0%) ^d	4 (21%)	0 (0%)	0 (0%)
Vision specific role difficulties	19 (41%)	5 (19%)	1 (14%)	1 (5%)	0 (0%)	0 (0%)
Vision specific dependency	16 (35%)	8 (30%)	1 (14%)	2 (11%)	0 (0%)	0 (0%)
Driving	18 (39%)	5 (19%)	2 (29%)	2 (11%)	0 (0%)	0 (0%)
Color vision	7 (15%)	2 (7%)	3 (43%)	1 (6%)	0 (0%)	0 (0%)
Peripheral vision	14 (30%)	4 (15%)	2 (29%)	1 (6%)	0 (0%)	0 (0%)

Proportions were compared on AS-20 and VFQ-25 composite scores and also between successes, partial successes and failures.

AS-20 = Adult Strabismus-20; VFQ-25 = National Eye Institute Visual Function Questionnaire-25

^a P≤0.01 significant difference between success and failure

^b P=0.04 significant difference between success and partial success

^c P=0.01 significant difference between partial success and failure

^d P=0.005 significant difference between AS-20 and VFQ-25 composite scores.

Table 4
Effect size statistic calculated for AS-20 and VFQ-25 in adults undergoing strabismus surgery.

	Effect size: Effect size (Mean change / Pre-op standard deviation)					
	Diplopia (n=80)		No diplopia (n=26)			
AS-20 questionnaire	Success (n=46)	Partial success (n=27)	Failure (n=7)	Success (n=19)	Partial success (n=5)	Failure (n=2)
AS-20 composite score	1.27 (25.3/19.8)	0.88 (17.5/19.9)	0.47 (8.0/17.1)	1.43 (21.8/15.3)	0.99 (12.9/13.1)	-0.18 (-3.1/17.7)
Psychosocial subscale	0.67 (17.3/25.5)	0.48 (11.3/23.6)	0.20 (5.0/25.2)	1.44 (30.0/20.9)	0.59 (16.4/27.7)	0.28 (2.5/8.8)
Function subscale	1.56 (33.3/21.3)	1.02 (21.0/20.6)	0.64 (11.1/17.2)	0.70 (13.7/19.6)	0.61 (9.5/15.5)	-0.20 (-8.8/44.2)
VFQ-25 questionnaire						
VFQ-25 composite score	1.14 (20.5/18.0)	0.69 (11.9/17.2)	0.54 (8.4/15.7)	0.57 (6.6/11.6)	0.41 (1.7/4.1)	-3.33 (-15.4/4.6)
General health	0.09 (2.2/23.2)	-0.04 (-0.9/26.3)	0.00 (0.0/23.8)	0.22 (5.3/24.0)	0.37 (5.0/13.7)	0.35 (25.0/70.7)
General vision	1.13 (23.5/20.8)	0.69 (12.6/18.3)	0.53 (11.4/21.4)	0.48 (7.4/15.4)	NA (12.0/0.0)	0.0 (0.0/14.1)
Ocular pain	0.68 (15.5/22.7)	0.28 (6.5/23.2)	0.11 (1.8/16.8)	0.22 (5.3/23.8)	0.0 (0.0/14.3)	-2.12 (-18.8/8.8)
Near activities	1.10 (23.7/21.5)	0.61 (15.7/25.8)	0.57 (10.1/17.6)	0.36 (5.3/14.5)	0.0 (0.0/4.2)	-0.71 (-16.7/23.6)
Distance activities	1.06 (21.6/20.4)	0.51 (12.8/25.3)	0.23 (4.8/20.8)	0.53 (8.8/16.4)	-0.67 (-5.0/7.5)	-0.24 (-4.2/17.7)
Vision specific social functioning	0.72 (14.9/20.7)	0.41 (7.9/19.2)	0.77 (10.7/13.9)	0.66 (10.5/15.9)	NA (0.0/0.0)	NA (-25.0/0.0)
Vision specific mental health	1.18 (34.0/28.7)	0.96 (22.2/23.2)	0.04 (0.9/20.5)	0.76 (14.5/19.1)	-0.80 (-5.0/6.3)	-0.80 (-17.7/22.1)
Vision specific role difficulties	0.97 (29.9/30.9)	0.58 (15.9/27.2)	0.56 (16.1/28.6)	0.32 (5.9/18.3)	0.45 (5.0/11.2)	NA (-37.5/0.0)
Vision specific dependency	0.64 (17.4/27.0)	0.60 (13.3/21.9)	0.18 (4.2/22.6)	0.29 (5.3/18.4)	0.45 (1.7/3.7)	NA (-16.7/0.0)
Driving	0.62 (18.7/30.2)	0.46 (10.3/22.3)	0.31 (6.9/22.6)	0.18 (3.9/22.0)	-0.15 (-1.7/10.9)	NA (-8.3/0.0)
Color vision	0.31 (4.9/16.0)	0.08 (1.0/11.8)	0.80 (10.7/13.4)	0.23 (1.3/5.7)	NA (0.0/0.0)	NA (-12.5/0.0)
Peripheral vision	0.80 (21.2/26.6)	0.46 (12.0/26.3)	0.62 (21.4/34.5)	0.18 (3.9/22.1)	0.45 (10.0/22.4)	-0.71 (-12.5/17.7)

Effect size is calculated by dividing mean pre- to post-operative change by pre-operative standard deviation. A higher effect size statistic indicates a more responsive measure. Results are shown for patients with and without diplopia, by success, partial success and failure. NA = not applicable / unable to calculate. AS-20 = Adult Strabismus-20; VFQ-25 = National Eye Institute Visual Function Questionnaire-25