



# HHS Public Access

Author manuscript

*Am J Ophthalmol.* Author manuscript; available in PMC 2017 April 01.

Published in final edited form as:

*Am J Ophthalmol.* 2016 April ; 164: 1–5. doi:10.1016/j.ajo.2015.12.029.

## Incorporating health-related quality of life into the assessment of outcome following strabismus surgery

Sarah R Hatt, David A Leske, Laura Liebermann, and Jonathan M Holmes

Department of Ophthalmology Mayo Clinic, Rochester, MN

### Abstract

**Purpose**—To evaluate changes in health-related quality of life (HRQOL) in adult strabismus patients classified as surgical failures by standard motor and diplopia criteria.

**Design**—Prospective cohort study evaluating outcomes

**Methods**—Adults undergoing strabismus surgery in a single clinical practice, with preoperative and 1-year postoperative Adult Strabismus-20 HRQOL questionnaires were included. Motor and diplopia criteria were applied to classify outcomes (success, partial success, or failure). For those classified failure, the medical record of the 1-year examination was reviewed to determine whether the patient reported subjective improvement. We evaluated improvement in HRQOL, defined as exceeding 95% limits of agreement on at least one of the four Adult Strabismus-20 domains. We compared proportions exceeding 95% limits of agreement in those reporting subjective improvement versus those who did not.

**Results**—40 (18%) of 227 patients were classified as failure by motor and diplopia criteria, with 39 of 40 able to exceed Adult Strabismus-20 95% limits of agreement. Overall, 21 (54%) of 39 showed improved HRQOL by exceeding 95% limits of agreement on at least one of the four Adult Strabismus-20 domains (54% vs predicted 10% by chance alone;  $P < 0.0001$ ). 25 (64%) reported subjective improvement, of whom 16 (64%) showed improved HRQOL exceeding 95% limits of agreement.

**Conclusions**—Many apparent surgical failures report subjective improvement, often reflected in improved HRQOL scores. We propose incorporating quantitative HRQOL criteria into the assessment of strabismus surgery outcomes, defining success as either meeting motor and diplopia criteria, or showing improvement in HRQOL beyond test-retest variability.

### Introduction

In adults, strabismus surgery outcomes are typically assessed by evaluating the angle of deviation.<sup>1-3</sup> In a previous study of outcomes we concluded that combining motor criteria and diplopia criteria provided a more representative assessment of surgical outcomes than

**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.

using either criterion on their own.<sup>3</sup> Combining motor and diplopia criteria raises the standard for success and reduces the likelihood that an unsatisfactory outcome would be misclassified as “success.” Nevertheless, combining criteria also means that patients experiencing significant improvement in motor alignment alone, and meeting criteria for motor success, could be classified as overall failures based on the presence of diplopia. In addition, many patients with profound ocular dysmotility preoperatively experience significant improvement in alignment and diplopia following surgery, yet without meeting criteria for success. The aim of this study was to assess patients classified as surgical failures by motor and diplopia criteria for evidence of subjective improvement following surgery and to analyze HRQOL scores to determine whether any documented subjective improvement was reflected by changes in HRQOL.

## Methods

Institutional Review Board approval for the review of participant’s clinical and questionnaire data in this retrospective cohort study evaluating outcomes, was obtained prior to commencement of the study from the Institutional Review Board at Mayo Clinic, Rochester, Minnesota, USA. All procedures and data collection were conducted in a manner compliant with the Health Insurance Portability and Accountability Act. All research procedures adhered to the tenets of the Declaration of Helsinki.

We retrospectively identified adult patients undergoing strabismus surgery in a single clinical practice who had completed the Adult Strabismus-20 questionnaire preoperatively (window: 1-28 days) and 1-year postoperatively (window: 5 months to 2 years). Postoperative questionnaire data were taken from the examination nearest to 1 year following surgery. Patients with any type of diplopic and non-diplopic strabismus were included; no exclusions were made based on diagnosis. Patients were not included if they were unable to read or understand English, had severe cognitive impairment, or had undergone additional extraocular muscle surgery before the 1-year outcome. All patients self-completed the Adult Strabismus-20 questionnaire, typically while in the waiting area and before any clinical testing was performed.

### Adult Strabismus-20 Questionnaire

The Adult Strabismus-20 questionnaire is a strabismus-specific questionnaire developed to assess the effects of strabismus on HRQOL in adults.<sup>4-6</sup> In Rasch analysis of the Adult Strabismus-20<sup>7</sup> four distinct domains were identified: Self-Perception, Interactions, Reading Function and General Function (full questionnaire freely available at: [www.pedig.net](http://www.pedig.net), accessed 10-21-2015). Each of the four domains are scored independently using Rasch scoring methods, and converted to a 0 to 100 score (worst to best HRQOL) for easier interpretation (scoring lookup table freely available at: [www.pedig.net](http://www.pedig.net), accessed 10-21-2015). Adult Strabismus-20 scores were calculated for each patient, for each of the four domains, at both the preoperative and the 1-year postoperative examination.

## Clinical Examination

As part of the standard clinical examination, the angle of deviation was measured in prism diopters (PD) using the simultaneous prism and cover test and the prism and alternate cover test at distance (3 meters) and near (1/3 meter) fixation, in habitual refractive correction. If the patient had poor visual acuity which precluded accurate prism cover test measurements, Krimsky measurements were used. If prism correction was worn, and only in-prism measurements were available, the simultaneous prism and cover test in prism correction was added to the prism strength to represent the underlying manifest angle of deviation.

All patients completed a diplopia questionnaire<sup>8</sup> as part of their clinical examination, rating the frequency of any diplopia in various gaze positions (straight ahead distance, reading, upgaze downgaze, right gaze, left gaze, any other position) as noticed over the past week. For each gaze position the frequency of any diplopia was rated as either never, rarely, sometimes, often, or always. This previously reported questionnaire was scored using a data driven scoring algorithm<sup>8</sup> on a 0 (no diplopia) to 100 (constant diplopia) scale (questionnaire and scoring algorithm freely available at: [www.pedig.net](http://www.pedig.net), accessed 10-21-2015).

## Classification of Postoperative Outcomes

Postoperative outcomes were classified using previously described motor and diplopia criteria.<sup>3</sup> For motor criteria, simultaneous prism and cover test angle data were used in order to capture the manifest misalignment. Postoperative outcomes were classified as “failure” if one of the following criteria were met: 1) simultaneous prism and cover test was 15 PD or more (horizontal or vertical) at distance or near; 2) diplopia or visual confusion was present more than “sometimes” straight ahead at distance or for reading (unless atypical diplopia due to decompensated childhood strabismus was present preoperatively, in which case diplopia was allowed postoperatively); 3) the patient was wearing an occlusive patch or Bangerter foil. Outcomes were classified as partial success if simultaneous prism and cover test was  $\leq$  15 PD (horizontal and vertical) at distance and near, and diplopia or visual confusion was present never, rarely or sometimes. Correction of diplopia with prism was allowed for classification as partial success. Outcomes were classified as success if simultaneous prism and cover test was  $<$  10 PD (horizontal and vertical) at distance and near, and diplopia or visual confusion was present never or only rarely.

For patients whose 1-year postoperative outcome was classified as failure, the history section of the medical record was reviewed to determine whether the patient had experienced any improvement following surgery. For example, statements by the patient that double vision or alignment was improved compared with preoperatively (for specific quotations see results). This assessment was made independently, without knowledge of preoperative or postoperative Adult Strabismus-20 HRQOL scores.

## Analysis

Change in preoperative to postoperative Adult Strabismus-20 scores was calculated and median preoperative to postoperative scores were compared using the signed-rank test. Improvement in HRQOL for an individual subject was defined as change in score exceeding

the 95% limits of agreement on at least one for the four Adult Strabismus-20 domains. The 95% limits of agreement define the limits within which 95% of differences in score due to test-retest variability should lie. Therefore a change in score exceeding these limits is likely to represent a meaningful change. In our present study, the 95% limits of agreement were calculated using 1.96 SD of previously reported test-retest differences<sup>9</sup> to define the limits. The 95% limits of agreement for the four Adult Strabismus-20 domains, calculated from these data, were 30.37 points for Self-Perception, 19.32 for Interactions, 24.75 for Reading Function, and 27.66 for General Function. Any patient unable to improve on at least one Adult Strabismus-20 domain (preoperative scores too high) was excluded from further analysis.

To determine whether reported subjective improvement in the medical record was reflected by a measured improvement in HRQOL, we compared the proportion of patients showing improved HRQOL in those who reported subjective improvement and those who did not, using Fishers exact tests.

## Results

Two hundred and twenty-seven adults undergoing surgery and with preoperative and 1-year postoperative Adult Strabismus-20 and diplopia questionnaire data were identified. One hundred and forty-two (63%) were female and 96% reported their race as White. Median age at the one-year examination was 53 (range 18 to 88) years. Overall, 40 (18%) of 227 were classified as failures, 128 (56%) were successes and 59 (26%) were partial successes. Data from 4 (10%) of 40 patients has been reported previously in prior studies.<sup>6, 10</sup>

### Reasons for Surgical Failure

Across 40 failures 32 (80%) were diplopic preoperatively, seven had no diplopia preoperatively and one had atypical diplopia associated with loss of suppression. Preoperative strabismus types were: cranial nerve palsies (n=15), restrictive (n=11), idiopathic / childhood onset (n=10), sensory (n=3), and other neuro (n=1). For 34 (85%) of 40 failures the reason for surgical failure was diplopia, rated as “often” or “always” for straight ahead distance or for reading; the remaining 6 (15%) patients failed for an simultaneous prism and cover test of 15PD or more at distance or near. Subsequent to the 1-year follow-up examination, 19 (48%) of 40 failures went on to receive additional surgery (18 incisional surgery and one Botulinum toxin injection).

### Median Improvement in Adult Strabismus-20 Scores for Failures

On average, preoperative and postoperative Adult Strabismus-20 scores improved numerically for each of the four Adult Strabismus-20 domains (Figure). Nevertheless, statistically significant improvement occurred only for the General Function domain, where median postoperative scores were significantly improved compared with preoperative scores (P=0.003, Figure).

### Improvement in HRQOL Exceeding Test-Retest Reliability

Thirty-nine of 40 failures had low enough pre-operative AS-20 scores that they were able to exceed 95% limits of agreement on at least one Adult Strabismus-20 domain and therefore were included for further analysis of postoperative improvement. Overall, 37 of 39 were able to exceed the 95% limits of agreement on at least two domains, 22 on at least 3 domains and 14 were able to exceed on all four domains. Overall 21 (54%) of 39 showed improved HRQOL by exceeding 95% limits of agreement on at least one Adult Strabismus-20 domain. This proportion is far greater than would be expected by test-retest variability alone. By definition, when using the 95% limits of agreement to define real change, only 2.5% would be expected to improve by test-retest variability (an additional 2.5% would be expected to worsen). Therefore, considering the 4 Adult Strabismus-20 domains, no more than 10% would be expected to improve by exceeding the 95% limits of agreement on at least one Adult Strabismus-20 domain (observed difference 54% vs 10%;  $P < 0.0001$ )

The Adult Strabismus-20 domain with the highest proportion exceeding 95% limits of agreement was Interactions (50% of those able to exceed, actually did), compared with 31% on General Function, 30% on Self-Perception, and 24% on Reading Function.

### Subjective Improvement Based on Medical Record

Twenty-five (64%) of 39 reported subjective improvement at the 1-year examination, whereas 14 did not. Of the 25 reporting subjective improvement by medical history, 16 (64%) also showed improved HRQOL, whereas of the 14 who did not report subjective improvement, 5 (36%) showed improved HRQOL (Table). Despite a numerically greater proportion showing improved HRQOL in those who reported subjective improvement (64% vs 36%), differences did not reach statistical significance ( $P = 0.09$ , Table).

Some examples of subjective improvement despite being designated a surgical failure by our criteria, are as follows: Case #1: preoperatively large-angle, consecutive, sensory exotropia. Postoperatively 2 PD esotropia but failed for new diplopia due to loss of suppression – patient reported being “very happy with alignment.” Case #2: preoperatively traumatic aphakia with incomitant strabismus developing following scleral buckle repair; previous eye muscle surgery elsewhere. Postoperatively failed for persistent binocular diplopia but patient reported being “100 times better than before surgery.” Case #3: preoperatively large convergence insufficiency type intermittent exotropia with diplopia. Postoperatively orthotropic at distance with residual intermittent exotropia with diplopia at near, patient reported that “surgery was a success.”

### Discussion

Of adult strabismus patients designated surgical failures by standard motor and diplopia criteria, we found that a large proportion (more than half) report subjective improvement and had measurable gains in HRQOL scores. This proportion was far greater than would be expected by test-retest variability alone. The most common domain of HRQOL improvement was in social interactions, consistent with the motor alignment success achieved for most of these patients (despite failure on diplopia criteria). Such gains from

strabismus surgery are important to recognize, since by clinical criteria these patients might be deemed unsuccessfully treated.

The standard for evaluating outcomes of strabismus surgery in adults has traditionally been motor alignment alone, with a commonly applied threshold for success being angle of deviation within 10 PD of orthotropia.<sup>1, 11, 12</sup> Diplopia criteria have been used less frequently and are more poorly standardized than motor criteria. In previous studies, diplopia success has been variously described in terms of complete resolution,<sup>13</sup> a specified size of the field of binocular single vision,<sup>14, 15</sup> or a minimal frequency on a patient-rated diplopia questionnaire.<sup>3</sup> Nevertheless, there are few studies reporting strabismus surgery outcomes in terms of changes in HRQOL. In a previous study by our group, we found a 60% success rate when success was defined by whether or not a patient exceeded 95% limits of agreement on at least one (of two) previous Adult Strabismus-20 domains (psychosocial and function).<sup>3</sup> Although clinical criteria, such as alignment and diplopia, remain the mainstay for assessing outcomes, consideration should be given to incorporating formal assessment of HRQOL as an alternative means of evaluating the effectiveness of surgical intervention. By limiting our outcome assessment to clinical measures only, we may overlook patient-reported improvement in strabismus-specific HRQOL and therefore misrepresent the wider-reaching benefits of surgical intervention.

Previous studies have shown overall improvement in both psychosocial- and function-related HRQOL in adults undergoing strabismus surgery<sup>5, 16-19</sup> but few differentiate between successful and unsuccessful surgery when evaluating HRQOL outcomes. We previously reported significantly greater gains in HRQOL for successfully treated patients compared with partial successes and surgical failures but of note, many of those designated surgical failure still showed a significant improvement in HRQOL from preoperative to postoperative examination.<sup>5, 6</sup> Based on the findings of this present study, such improvements in HRQOL may be explained by the fact that real improvements in HRQOL are occurring in the presence of reported subjective improvement, despite the patient not achieving surgical success by motor and diplopia criteria.

Formal assessment of HRQOL using, for example, the Adult Strabismus-20 questionnaire provides a validated, strabismus-specific patient-reported outcome measure for adults undergoing strabismus surgery. For an individual patient, real improvement in HRQOL is best assessed by whether a change in score exceeds normal testing variability, as defined by 95% limits of agreement. In a previous study<sup>3</sup> we found that exceeding the 95% limits of agreement was a high hurdle to clear, with some successfully aligned patients not exceeding 95% limits of agreement on the Adult Strabismus-20. Nevertheless, the fact that exceeding 95% limits of agreement is a challenge, when they are exceeded (as they were for some failures in this present study), this should be considered an important achievement.

It is possible that our finding of improved HRQOL in the context of surgical failure is attributable to a placebo effect of surgery on HRQOL unrelated to success or failure. Nevertheless, we have previously demonstrated construct validity of the Adult Strabismus-20, reporting that patients with more favorable clinical outcomes have greater improvement in HRQOL scores following surgery.<sup>5, 6</sup> Although it may be impossible to

determine whether the HRQOL improvement observed in the present study is at least in part the result of a placebo effect, it should be noted that HRQOL scores reported in this study are 1 year following surgery and we would expect any placebo effect would be minimal if there was truly an unfavorable clinical outcome.

It is possible that we found improved HRQOL in surgical failures because our definition of failure was too inclusive, and that some failures should in fact have been classified as a partial success or even a success. Nevertheless, most patients failed for persistent or new diplopia rated as present “often” or “always” for reading or for straight ahead distance, and it does not seem reasonable to designate such a high frequency of diplopia in key gaze positions as surgical success. Alternatively, it may be that patient interpretation of “often” or “always” is not consistent. For example, what one patient might consider “often,” another patient with the same severity of diplopia might consider “sometimes” or even “rarely.” Further refinement of the diplopia questionnaire, specifying an approximate frequency for each level of severity may enable better classification of outcomes.

Our study is not without limitations. The medical history was not always clear regarding the presence versus absence of subjective improvement and it is possible that some patients would have been classified differently had each patient been asked directly regarding improvement following surgery.

In summary, many adult strabismus patients designated surgical failure by standard motor and diplopia criteria, show measurable improvement in HRQOL. Quantitative assessment of HRQOL is an important component of outcome assessment in adults undergoing strabismus surgery, and improvements in HRQOL may indicate important treatment-related gains to the patient, even when surgery might be otherwise be considered a failure. Consideration should be given to redefining outcome criteria for adults undergoing strabismus surgery such that a patient be classified as “success” not only when meeting specific motor and diplopia criteria, but also when showing improvement in HRQOL that exceeds test retest variability.

## Acknowledgements

a. *Funding/Support:* This study was supported by National Institutes of Health Grant EY018810 (JMH), Research to Prevent Blindness, New York, New York (JMH as Olga Keith Weiss Scholar and an unrestricted grant to the Department of Ophthalmology, Mayo Clinic), and Mayo Foundation, Rochester, Minnesota.

b. *Financial Disclosures:* No authors have any financial/conflicting interests to disclose.

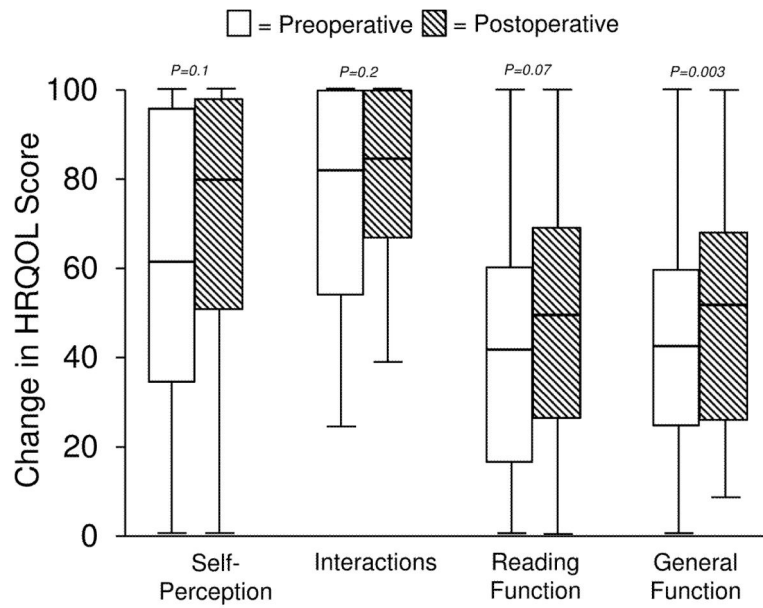
c. *Other Acknowledgments:* None.

## References

1. Mills MD, Coats DK, Donahue SP, Wheeler DT, American Academy of Ophthalmology. Strabismus surgery for adults: a report by the American Academy of Ophthalmology. *Ophthalmology*. 2004; 111(6):1255–1262. [PubMed: 15177982]
2. Kushner BJ. The efficacy of strabismus surgery in adults: a review for primary care physicians. *Postgrad Med J*. 2011; 87(1026):269–273. [PubMed: 21273365]
3. Hatt SR, Leske DA, Liebermann L, Holmes JM. Comparing outcome criteria performance in adult strabismus surgery. *Ophthalmology*. 2012; 119(9):1930–1936. [PubMed: 22541935]

4. Hatt SR, Leske DA, Bradley EA, Cole SR, Holmes JM. Development of a quality-of-life questionnaire for adults with strabismus. *Ophthalmology*. 2009; 116(1):139–144. [PubMed: 19019449]
5. Hatt SR, Leske DA, Holmes JM. Responsiveness of health-related quality of life questionnaires in adults undergoing strabismus surgery. *Ophthalmology*. 2010; 117(12):2322–2328. [PubMed: 20832120]
6. Hatt SR, Leske DA, Liebermann L, Holmes JM. Changes in health-related quality of life 1 year following strabismus surgery. *Am J Ophthalmol*. 2012; 153(4):614–619. [PubMed: 22285013]
7. Leske DA, Hatt SR, Liebermann L, Holmes JM. Evaluation of the Adult Strabismus-20 (AS-20) Questionnaire using Rasch analysis. *Invest Ophthalmol Vis Sci*. 2012; 53(6):2630–2639. [PubMed: 22447864]
8. Holmes JM, Liebermann L, Hatt SR, Smith SJ, Leske DA. Quantifying diplopia with a questionnaire. *Ophthalmology*. 2013; 120(7):1492–1496. [PubMed: 23531348]
9. Leske DA, Hatt SR, Holmes JM. Test-retest reliability of health-related quality of life questionnaires in adults with strabismus. *Am J Ophthalmol*. 2010; 149(4):672–676. [PubMed: 20138603]
10. Hatt SR, Leske DA, Liebermann L, Philbrick KL, Holmes JM. Depressive symptoms associated with poor health-related quality of life in adults with strabismus. *Ophthalmology*. 2014; 121(10):2070–2071. [PubMed: 24908202]
11. Carruthers JD, Kennedy RA, Bagaric D. Botulinum vs adjustable suture surgery in the treatment of horizontal misalignment in adult patients lacking fusion. *Arch Ophthalmol*. 1990; 108(10):1432–1435. [PubMed: 2222277]
12. Zhang MS, Hutchinson AK, Drack AV, Cleveland J, Lambert SR. Improved ocular alignment with adjustable sutures in adults undergoing strabismus surgery. *Ophthalmology*. 2012; 119(2):396–402. [PubMed: 22036633]
13. Beauchamp GR, Black BC, Coats DK, et al. The management of strabismus in adults--I. Clinical characteristics and treatment. *J AAPOS*. 2003; 7(4):233–240. [PubMed: 12917608]
14. Hertle RW. Clinical characteristics of surgically treated adult strabismus. *J Pediatr Ophthalmol Strabismus*. 1998; 35:138–145. [PubMed: 9627873]
15. Nassar MM, Dickinson AJ, Neoh C, et al. Parameters predicting outcomes of strabismus surgery in the management of Graves' ophthalmopathy. *J AAPOS*. 2009; 13(3):236–240. [PubMed: 19395292]
16. Glasman P, Cheeseman R, Wong V, Young J, Durnian JM. Improvement in patients' quality-of-life following strabismus surgery: evaluation of postoperative outcomes using the Adult Strabismus 20 (AS-20) score. *Eye (Lond)*. 2013; 27(11):1249–1253. [PubMed: 23928880]
17. Jackson S, Harrad RA, Morris M, Rumsey N. The psychosocial benefits of corrective surgery for adults with strabismus. *Br J Ophthalmol*. 2006; 90(7):883–888. [PubMed: 16782950]
18. Burke JP, Leach CM, Davis H. Psychosocial implications of strabismus surgery in adults. *J Pediatr Ophthalmol Strabismus*. 1997; 34(3):159–164. [PubMed: 9168420]
19. Beauchamp GR, Black BC, Coats DK, et al. The management of strabismus in adults--III. The effects on disability. *J AAPOS*. 2005; 9(5):455–459. [PubMed: 16213395]





**Figure.** Box plots of preoperative and postoperative Adult Strabismus-20 scores in 40 adults who apparently failed strabismus surgery by standard motor and diplopia criteria. Boxes represent first quartile, median, and third quartile values; whiskers represent extreme values. *P* value represents difference between preoperative and postoperative examination scores.

**Table**

Subjective improvement and improvement in health-related quality of life in adults designated strabismus surgery “failures” by motor and diplopia criteria.

	<b>Subjective improvement (N=25)</b>	<b>No subjective improvement (N=14)</b>	<b>Difference between groups</b>
HRQOL improved (N=21)	16 (64%)	5 (36%)	P=0.09
HRQOL not improved (N=18)	9 (36%)	9 (64%)	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript