Published in final edited form as:

Am J Ophthalmol. 2011 November ; 152(5): 872–876. doi:10.1016/j.ajo.2011.05.007.

Improved Assessment of Control in Intermittent Exotropia Using Multiple Measures

Sarah R. Hatt, Laura Liebermann, David A. Leske, Brian G. Mohny, and Jonathan M. Holmes

Department of Ophthalmology, Mayo Clinic, Rochester, MN

Abstract

Purpose—To develop and validate an improved measure of control in intermittent exotropia (XT).

Design—Prospective, non-interventional case series.

Methods—Twelve children with intermittent XT were evaluated during 4 sessions (2 hours apart) over a day, on 2 separate days (8 sessions per child). Control was standardized using a scoring system and quantified 3 times during each examination. Overall control for a day was calculated as the mean of all 12 measures. Single measures of control and the mean of 2 (double) and 3 (triple) measures over the examination were compared with the respective day mean, and first day measures were compared to the second day mean.

Results—At distance, 17% (49/287, 95% CI 13% to 22%) of single measures differed from the day mean whereas only 8% (16/191, 95% CI 5% to 13%) of double measures and 5% (5/95, 95% CI 2% to 12%) of triple measures differed. Comparing day 1 measures to overall mean for day 2, 17% (24/143, 95% CI 11% to 24%) of single measures and 17% (22/130, 95% CI 11% to 24%) of double measures differed by more than 1 level, whereas 11% (5/47, 95% CI 4% to 23%) of triple measures differed.

Conclusions—The mean of 3 assessments of control during a clinic examination better represents overall control than a single measure.

Introduction

In intermittent exotropia (XT), severity is often judged by assessing a patient's ability to control the underlying exodeviation. Assessment of control is most often based on observations of spontaneous motor fusion, or lack of spontaneous motor fusion, either by the clinician or the parent.^{1–5} Some authors have stated that “poor control” should be an indication for surgery, but “poor control” was not well defined.^{6–8} Recently, attempts have been made to quantify control using control scales.^{1–5} Nevertheless, these control scores rely on a single assessment during the clinical examination and in a previous study⁹ we found considerable variability when comparing isolated measures of control, including change

© 2011 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.

Conflict of Interest: No conflicting relationship exists for any author

from spontaneous tropia to phoria (and vice versa). In our previous study, changes in control occurred both over the course of one day (approximately 2 hours between assessments), and also from one minute to the next when no intervening testing or dissociation had occurred. Such variability makes it difficult to reliably distinguish between a patient with consistently poor control and a patient whose control assessment happens to coincide with a rare episode of poor control. In a previous report, we postulated that multiple measures of control might better represent overall control for an individual patient.⁹ In the present study, we compared 1, 2 and 3 assessments of control during a single office examination session to an overall assessment of control during the day.

Methods

Patients

Parents/legal guardians of eligible children were contacted about the study and 12 children (median age 7, range 6 to 13 years) were prospectively enrolled. Children were considered eligible if they had basic, pseudo or true types of intermittent XT; convergence insufficiency type intermittent exotropia (near angle more than 10 prism diopters [pd] greater than distance), sensory exotropia, paralytic exotropia or co-existing developmental delay were excluded. Due to the testing burden, the minimum age for inclusion was set at 5 years old. The median visual acuity was 20/20, range 20/15 to 20/30 in the right eye and median 20/22, range 20/20 to 20/30 in the left eye. No patient had any inter-ocular visual acuity difference of more than 2 LogMAR lines. Nine (75%) were female, consistent with previously reported female preponderance in intermittent XT,¹⁰ and race was reported as 'white' for 10 (83%). The median angle of deviation by prism and alternate cover test (PACT) was 20 (range 14 to 30) prism diopters (pd) at distance (3 meters) and 12 (range 6 to 25) pd at near.

Control assessment

Assessment of control was standardized using a previously described 0- to 5-point control scale⁵ scored separately for distance fixation (3 meters) and then for near fixation (1/3 meter). Using this control scale, any spontaneous exotropia was graded during a 30-second period of observation (constant exotropia = score 5, exotropia more than 50% of the time = score 4, exotropia less than 50% of the time = score 3).⁵ In the absence of any spontaneous exotropia during the 30-second observation period, exodeviation recovery was rated as the worst of three, 10-second dissociations (more than 5 seconds to recovery = score 2, 1–5 seconds to recovery = score 1, less than 1 second to recovery = score 0).⁵

Examinations

Children were evaluated at 4 separate examinations over the course of a day, with at least 2 hours between each examination. Assessments were repeated in the same manner on a second day (median 21 days, range 7 to 41 days later), yielding a total of 8 examinations per child. It was not possible to standardize the amount of time between first and second days because families often had many other conflicting commitments. All examination procedures were performed by a certified orthoptist. At each examination, testing was performed according to a pre-determined testing protocol. Control was assessed 3 times during each examination; at the start of the examination, after the PACT, and at the end of the examination, yielding a total of 12 control assessments over the day. The median duration of each examination was 21 minutes (range, 10 to 55 minutes).

Analysis

For the purposes of this study, an overall control score for the day was calculated for each patient as the mean of all 12 control scores (day mean). Therefore, over the 2 days there

were a total of 24 control scores per child and 2 day mean scores per child. One child was unable to complete one assessment of control at one examination due to waning cooperation and therefore the total number of control assessments over the 2 days was 287 out of a potential 288. Single control scores (n=287) were compared to the respective day mean, calculating proportions differing by more than 1 level. Double control scores were calculated as: 1) the mean of first and second control scores during a single examination session and 2) the mean of first and third control scores during a single examination session. These combinations for double scores were used in order to reflect clinical scenarios where control might be measured at the beginning and end of the examination or at the beginning and middle of the examination. Triple control scores were calculated as the mean of the 3 control assessments across a single examination session. Double control scores (2 per examination; n=191) and triple control scores (1 per examination; n=95) were then compared to the respective day mean, calculating proportions differing by more than 1 level. Distance and near control were analyzed separately.

In order to evaluate whether single, double or triple control scores at one examination session represent an individual patient, we compared single (n=143), double (n=130) and triple (n=47) scores on day 1 to the overall mean (of all 12 control scores) for day 2, calculating proportions differing by more than 1 level. Day 1 and day 2 mean control scores (mean of 12 measures across the day) were compared using Intraclass Correlation Coefficients (ICC).

Results

Variability over the day and over the exam

Consistent with our previous study of variability of single control scores,⁹ 11 of 12 (92%) patients showed a change (> 1 level) over the day at either distance or near on either day 1 or day 2. Seven of 12 (58%) patients changed at distance or near on both day 1 and day 2. Consistent with our previous report of variability over an examination,⁹ 8 of 12 (67%) patients showed a change in single measures of control on at least one examination at distance and 6 of 12 (50%) at near.

Single, double and triple measures of control at distance

For distance, 17% (49 of 287, 95% CI 13% to 22%) of single measures and 8% (16 of 191, 95% CI 5% to 13%) of double measures differed by more than 1 level from the respective day mean (Table 1), whereas only 5% (5 of 95, 95% CI 2% to 12%) of triple measures differed (Table 1).

Single, double and triple measures of control at near

For near, 12% (34 of 287, 95% CI 8% to 16%) of single measures and 4% (8 of 191, 95% CI 2% to 8%) of double measures differed by more than 1 level from the respective day mean (Table 1), whereas only 1% (1 of 95, 95% CI 0% to 6%) of triple measures differed (Table 1).

Correlation of mean control over day one and day two

At distance, day 1 and day 2 means were highly correlated (ICC=0.88, 95% CI 0.64 to 0.96, Figure 1 top). At near, day 1 and day 2 means were moderately correlated (ICC= 0.58, 95% CI 0.06 to 0.85, Figure 1 bottom).

Comparing composite day 1 control scores to day 2 mean score

To evaluate whether single, double or triple control scores at one examination represent an individual patient, single, double and triple scores on day 1 were compared to the overall mean for day 2. At distance 17% (24 of 143, 95% CI 11% to 24%) of single measures and 17% (22 of 130, 95% CI 11% to 24%) of double measures differed by more than 1 level (Table 2), whereas only 11% (5 of 47, 95% CI 4% to 23%) of triple measures differed (Table 2). At near, 32% (46 of 143, 95% CI 25% to 40%) of single measures differed by more than 1 level whereas 22% (29 of 130, 95% CI 15% to 30%) of double measures and 21% (10 of 47, 95% CI 11% to 36%) of triple measures differed (Table 2).

Discussion

In the present study we confirmed our previous findings,⁹ that exodeviation control varies over a day, and over an examination, in most children with intermittent XT. We have now established that a triple control measure (the mean of 3 measures during an examination) is closer to mean control over the day than a single control measure, both over the same day and for a subsequent day. Therefore, a triple control score better represents overall control in an individual patient.

Previous studies have reported different approaches to rating control of intermittent XT during a clinic examination.^{1–5} Nevertheless there are few reliability data on these methods and the time for control assessment is typically not standardized. Some approaches incorporate parental reports of frequency in an attempt to better capture a patient's overall control,³ but the reliability of parental reports is unknown.

We are not aware of previous data examining the use of multiple measures over an examination to improve the reliability of control assessment in intermittent exotropia. We previously reported variability in single measures of control from minute to minute and over one day in a different cohort of children with intermittent XT⁹ and highlighted the need for a control measure that better captures variability over the examination and over the day. The findings of this present study suggest that, using the control scale utilized in this study, the mean of 3 control measures (triple control score), provides a better summary of control than single or double measures.

Averaging repeated measurements will, by definition, reduce any variability and provide a value closer to the patient's true status. It is therefore not particularly surprising that we found the mean of 2 (double) and 3 (triple) measures over an examination to be closer to the overall mean for the day than single measures. Although a double control score provides valuable additional information compared to a single control measure, the triple score (mean of 3 measures) was significantly different from a single measure and closer to the overall day mean in most patients. We found that 3 control assessments could be easily performed over the course of a regular clinic examination in most children with intermittent XT, but more than 3 assessments would likely prove onerous.

We further assessed the role of single, double and triple measures of control during an examination, in representing an individual patient with intermittent XT, by comparing single, double and triple measures on day 1 to overall mean control on day 2. This method is particularly appropriate because we found that the mean of 12 measures on day 1 was correlated to the mean of 12 measures on day 2, despite variability between measures throughout each day. We again found better agreement for triple measures of control with the overall day mean, than for double or single measures, especially at distance fixation.

Variable control over one day in children with intermittent XT renders isolated measures of control inadequate for representing severity in an individual patient. Using single measures of control to determine severity could then result in misclassification of severity at a given examination and consequently misinterpretation of apparent change in severity from one visit to the next. Using a triple control score during an examination appears to provide an improved summary measure compared to a single control score, better capturing overall severity and therefore more closely representing control status for an individual patient at a given point in time. In future studies it will be interesting to explore change in triple measures over time and examine correlations with other parameters such as angle of deviation, stereoacuity and motor fusion. It is unclear whether deterioration in control should be used as the sole determinant for surgery in intermittent XT, but in order to study this question, robust methods of assessing control are needed and our proposal of a triple assessment of control may provide such a method.

Due to the burden of testing and the demands on patient's time, we were only able to recruit a small number of patients to this study and our conclusions would have been strengthened if we had been able to study a larger cohort. In addition, it would have been preferable to standardize the time between the first and second day. We also limited the maximum number of control measures over an examination to 3, but we have found it impractical to perform more than 3 formal assessments of control during a single examination.

Most children with intermittent XT show variable control of their exodeviation and the mean of triple control scores during an examination better captures overall control (as measured over one day) compared with a single measure. Triple measures of control provide a more robust method of representing control severity for an individual patient.

Acknowledgments

- a. *Funding / Support:* Supported by National Institutes of Health Grants EY015799 and EY018810 (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.
- b. *Financial Disclosures:* Grant support by National Institutes of Health Grants EY015799 and EY018810 (JMH), and Research to Prevent Blindness, New York, NY (JMH as Olga Keith Weiss Scholar and an unrestricted grant to the Department of Ophthalmology, Mayo Clinic)
- c. *Contributions to Authors:* design and conduct of the study (SRH, LL, DAL, BGM, JMH); collection of data (SRH, LL, JMH), management, analysis, and interpretation of the data (SRH, DAL, LL, JMH, BGM); preparation, review and approval of the manuscript (SRH, DAL, LL, BGM, JMH).
- d. *Statement about Conformity with Author Information:* Institutional Review Board/Ethics Committee approval was obtained for this study from Mayo Clinic, Rochester MN. All experiments and data collection were conducted in a manner compliant with the Health Insurance Portability and Accountability Act. Informed consent was obtained from all participants and informed assent for participants aged 8 to 12 years old. All research procedures adhered to the tenets of the Declaration of Helsinki.
- e. *Other Acknowledgments:* None

References

1. Stathacopoulos RA, Rosenbaum AL, Zanon D, et al. Distance stereoacuity. Assessing control in intermittent exotropia. *Ophthalmology*. 1993; 100(4):495–500. [PubMed: 8479706]
2. Petrunak JL, Rao R. The evaluation of office control in intermittent exotropia: a systematic approach. *Am Orthopt J*. 2003; 53:98–104. [PubMed: 21149078]
3. Haggerty H, Richardson S, Hrisos S, Strong NP, Clarke MP. The Newcastle Control Score: a new method of grading the severity of intermittent distance exotropia. *Br J Ophthalmol*. 2004; 88(2): 233–235. [PubMed: 14736781]

4. Chia A, Seenyen L, Long QB. A retrospective review of 287 consecutive children in Singapore presenting with intermittent exotropia. *J AAPOS*. 2005; 9(3):257–263. [PubMed: 15956946]
5. Mohny BG, Holmes JM. An office-based scale for assessing control in intermittent exotropia. *Strabismus*. 2006; 14(3):147–150. [PubMed: 16950743]
6. von Noorden, GK.; Campos, EC. Exodeviations. In: Lampert, R.; Cox, K.; Burke, D., editors. *Binocular Vision and Ocular Motility. Theory and Management of Strabismus*. 6th ed.. St. Louis: Mosby; 2002. p. 356-376.
7. Santiago, AP.; Ing, MR.; Kushner, BJ.; Rosenbaum, AL. Intermittent exotropia. In: Rosenbaum, AL.; Santiago, AP., editors. *Clinical Strabismus Management Principles and Surgical Techniques*. Philadelphia: W.B: Saunders Company; 1999. p. 163-175.
8. Wright, K. Exotropia. In: Wright, K., editor. *Pediatric ophthalmology and strabismus*. St Louis: Mosby Year Book; 1995. p. 195-202.
9. Hatt SR, Mohny BG, Leske DA, Holmes JM. Variability of control in intermittent exotropia. *Ophthalmology*. 2008; 115(2):371–376. [PubMed: 17629562]
10. Nusz KJ, Mohny BG, Diehl NN. Female predominance in intermittent exotropia. *Am J Ophthalmol*. 2005; 140(3):546–547. [PubMed: 16139014]

Figure 1.

Correlation between Day 1 and Day 2 mean control scores in children with intermittent exotropia. For each child, the day mean was calculated as the mean of 12 measures of control across the day. Mean day scores were highly correlated at distance (Top: ICC=0.88, 95% CI 0.64 to 0.96) and moderately correlated at near (Bottom: ICC= 0.58, 95% CI 0.06 to 0.85). Dotted line represents unity line.

Table 1

Agreement of day 1 and day 2 single, double (mean of 2) and triple (mean of 3) control scores with respective day score (mean of all 12 measures across the day), in 12 children with intermittent exotropia. Table shows the proportion of single, double and triple scores differing more than 1 level from the respective day score.

	Single control scores	Mean of 2 (double) control scores	Mean of 3 (triple) control scores
Distance	17% (49/287, 95% CI 13% to 22%)	8% (16/191, 95 CI 5% to 13%)	5% (5/95, 95% CI 2% to 12%)
Near	12% (34/287, 95% CI 8% to 16%)	4% (8/191, 95% CI 2% to 8%)	1% (1/95, 95% CI 0% to 6%)

Table 2

Agreement of day 1 single, double (mean of 2) and triple (mean of 3) control scores with day 2 overall score (mean of all 12 measures across the day), in 12 children with intermittent exotropia. Table shows the proportion of single, double and triple scores on day 1 differing more than 1 level from the mean day 2 score.

	Single control scores	Mean of 2 (double) control scores	Mean of 3 (triple) control scores
Distance	17% (24/143, 95% CI 11% to 24%)	17%, (22/130, 95% CI 11% to 24%)	11% (5/47, 95% CI 4% to 23%)
Near	32% (46/143, 95% CI 25% to 40%)	22% (29/130, 95% CI 15% to 30%)	21%, (10/47, 95% CI 11% to 36%)