



Published in final edited form as:

Ophthalmology. 2008 February ; 115(2): 371–376.e2. doi:10.1016/j.ophtha.2007.03.084.

Variability of control in intermittent exotropia

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

From Department of Ophthalmology, Mayo Clinic College of Medicine, Rochester, MN

Abstract

Purpose—Assessing control is thought to be important in the management of intermittent exotropia (IXT), including the decision to perform surgery. The purpose of this study was to assess the presence and degree of any change in control occurring over the course of 1 day using a previously described 6-point clinic control scale, and to evaluate inter-observer and minute-to-minute variability.

Design—Prospective case series.

Participants—25 patients with intermittent exotropia.

Methods—Inter-observer agreement was determined in 17 patients by comparing control scores assessed simultaneously by 2 observers (kappa test). Minute-to-minute variability was observed in the same 17 patients by assessing control twice within 5 minutes. Variability over one day was assessed in 5 of these patients plus 8 additional patients (n=13) by comparing 3 or 4 assessments at least 2 hours apart.

Main outcome measure—Control of IXT measured using a 6-point clinic control scale.

Results—Inter-observer agreement was high ($k=0.94$ for distance and $k=0.95$ for near fixation). Disagreements were no more than one level on the control scale; therefore, for further analysis, change in control was defined as 2 or more levels. For minute-to-minute variability, 4 (24%) of the 17 patients tested twice within 5 minutes showed a change in control: 1 (6%) changed from tropia to phoria at distance, 3 (18%) from phoria to tropia at near. Of the 13 patients assessed over 1 day, 6 (46%) showed change in control, 2 at distance fixation only, 3 at near only and 1 at near and distance.

Conclusions—Control of IXT can vary throughout the day, even within minutes, including from phoric to spontaneously tropic, and vice versa. The worst level of control was not always later in the day. This suggests that an isolated assessment of control may not categorize severity of IXT in an individual patient.

Intermittent exotropia (IXT) is a common form of childhood strabismus^{1, 2} characterized by intermittent divergent misalignment that is often greater at distance fixation. Severity and progression of IXT are currently assessed by determining angle of deviation, stereoacuity, ability to control the deviation and parental reports of frequency. ‘Control’ may be defined as an estimate of the proportion of time the deviation is manifest and the ease of re-establishing fusion after dissociation. Although the natural history and indications for surgical intervention are poorly defined,³ it is commonly held that poor control of the exodeviation is a sign of deterioration.⁴ On this basis, surgery has been recommended by some authors when the exotropia occurs more than 50% of waking hours or when there is a gradual loss of fusional

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.

No conflicting relationships exist

This article contains online-only material. The following should appear online-only: Table 2 and Table 3.

control as evidenced by an increasing frequency of the exotropia.⁴⁻⁶ Others have suggested that early surgery results in superior sensory outcomes^{7, 8} but the timing of surgery for IXT has not been subjected to a randomized clinical trial and prior to such a trial, outcome measures such as control need to be rigorously studied.

Until recently, control has not been quantified on an ordinal scale but has been documented qualitatively by describing recovery on the cover test and documenting parental reports of frequency. Control scales⁹⁻¹³ allow quantification of control, but it remains unclear whether these measures reliably represent severity. The purpose of this study was to assess changes in control occurring over the course of a day and a period of several minutes using a previously described 6-point control scale¹⁰ in patients with intermittent exotropia.

Methods

Institutional Review Board/Ethics Committee approval was obtained for this study. All experiments and data collection were conducted in a manner compliant with the Health Insurance Portability and Accountability Act.

Patients

We prospectively enrolled 25 patients with IXT for three concurrent parts of this study: inter-observer variability, minute-to-minute variability, and intraday variability. Patients were excluded if they had convergence insufficiency type exotropia (near angle ≥ 10 prism diopters greater than distance) or co-existing ocular pathology. Patients were assessed in their habitual refractive correction. No patient had amblyopia defined as ≥ 0.2 LogMAR (Logarithm of the Minimum Angle of Resolution) inter-ocular difference and ≥ 0.3 LogMAR in one eye. If too young for optotype testing amblyopia was evaluated on the basis of fixation preference.

Assessment of control

For each patient, control of the IXT was assessed using a previously described control scale¹⁰ rating phoria from 0-2 and tropia from 3-5 for both distance (3 m) and near (1/3 m) fixation (Table 1).^{10, 14}

Inter-observer variability

Our underlying hypothesis was that control varies within short periods of time. For this reason we could not rely on sequential assessments to evaluate inter-observer variability. We therefore assessed inter-observer variability by simultaneous observation, one investigator performing the test and a second observing. In this way, distance and near control were independently rated by 2 investigators in 17 patients with IXT, 13 children and 4 adults (overall median age 8 years; range 1 to 33 years). In the first examination / observation pair (time one), investigator one assessed control for each patient while investigator two observed. Control score values were recorded independently, without knowledge of the other investigator's assessment. The paired examination was repeated within 5 minutes (time two) when investigator two assessed control and investigator one observed. For analysis we therefore used 34 pairs of near scores (17 at time one and 17 at time two) and 34 pairs of distance scores (17 at time one and 17 at time two), giving a total of 68 simultaneous pairs. Agreement between examiners was determined using the Kappa test. The results from this part of the study were used to establish the threshold for determining real change in control for the assessments performed over several minutes and throughout the day.

Minute-to-minute variability

Data for evaluation of minute-to-minute variability were derived from the same patients and same examinations as the inter-observer part of the study. For each patient, control ratings on the first assessment (time one) were compared to those on the second assessment (time two), whether the investigator was performing or observing the assessment, i.e., investigator one examining at time one and observing at time two and investigator two observing at time one and examining at time two. A total of 34 assessments for investigator one (17 near and 17 distance) and 34 for investigator two (17 near and 17 distance) were analyzed: near and distance control scores were compared separately.

Intraday variability

Thirteen children (median age 8 years; range 1 to 13 years) with IXT were assessed by a single investigator 3 or 4 times throughout the day (5 of the 13 patients were also independently recruited to the minute- to minute study but with 6 weeks to 2 years between studies). Assessments over the day were scheduled during the following time intervals: 08:00 – 10:30; 10:31 -13:00; 13:01 - 15:30 and 15:31 - 18:00. Near and distance control and alternating prism cover test were measured at each time point. Patients were categorized as either showing change or no change in control.

Analysis of control and angle during intraday study

The median control score and median angle of deviation (prism diopters) were calculated for near and distance fixation for each of the 13 patients in the intraday part of the study. Correlations between median near control score and near angle, and median distance control score and distance angle were evaluated using the Spearman test.

Results

Inter-observer variability

There was almost perfect agreement between examiners for control scores at distance fixation ($k=0.94$) (Figure 1) and at near fixation ($k=0.95$) (Figure 2). Disagreements occurred in only 5 (7%) of 68 assessments and differed no more than one level on the scale. There were no instances when one examiner recorded a phoria (0-2 on scale) while the other recorded a tropia (3-5 on scale). A change of 2 or more levels on the scale therefore defined the threshold for a real change in control for subsequent parts of the study.

Minute-to-minute variability

As rated by investigator one, 4 (24%) of 17 patients showed a change in control (2 or more levels on the control scale) between time one and two: 1 (6%) of 17 changed from tropia to phoria at distance, 3 (18%) of 17 changed from phoria to tropia for near (Table 2; available at <http://aaojournal.org>). As rated by investigator two, the same 4 patients showed the same pattern of change between times one and two (Table 2; available at <http://aaojournal.org>). The magnitude of change was different in one case: patient 12 (Table 2; available at <http://aaojournal.org>) was recorded as changing by 2 levels on the scale by examiner one and 3 levels on the scale by examiner two.

Intraday variability

Six (46%) of the 13 patients showed a change in score (2 or more levels on the control scale): 2/6 showed a change at distance only, 3/6 at near only and 1/6 at both distance and near (Table 3; available at <http://aaojournal.org>). All 3 patients who changed at distance fixation, changed from spontaneous tropia (score 3-5) to phoria (score 0-2) (Figure 3) at one time point during the day. Two of the 4 who changed at near, went from spontaneous tropia to phoria (Figure 4)

at one time point during the day, although for one of these, the change from tropia to phoria was a change of only one point on the scale. The maximum change was 3 levels, occurring in 4 patients; 1 for near and 3 for distance (Table 3; available at <http://aaajournal.org>). In all 6 cases the last assessment of the day did not show the worst level of control (Table 3; available at <http://aaajournal.org>, Figures 3 and 4).

Analysis of control and angle during intraday study

In the 13 patients studied over one day, there was an overall correlation between median near angle of deviation and median near control score ($r=0.7$; $p=0.02$) but not between median distance angle of deviation and median distance control score ($r=-0.2$; $p=0.5$).

Comparison of patients with and without change during intraday study

Using our previously defined criteria for change in control score (at least 2 levels on the scale), 6 (46%) of 13 patients changed at distance or near during the day and 7 (54%) of 13 did not. Regarding patient age, there was no difference between those who changed and those who did not (mean 7.3 ± 4.1 years vs 7.3 ± 2.7 years; $p=0.98$). Regarding median control score, those who changed, whether for distance or near, appeared to show worse control at distance than those who did not (median 2.75 vs 1, $p=0.04$). Nevertheless, there was no difference in the median control at near between those who changed and those who did not. (median 1 vs 1; $p=0.2$).

Analyzing by median angle during the day, there was no significant difference between those showing change in control and those who did not, for either distance fixation (mean 26 ± 5 vs 30 ± 10 prism diopters; $p=0.4$) or near fixation (mean 25 ± 11 vs 20 ± 12 prism diopters; $p=0.5$).

Discussion

Based on assessment of inter-observer variability, we found that grading control of intermittent exotropia using the clinic control scale¹⁰ provides a reliable means of quantifying the level of control at any given point in time. Nevertheless, in a large proportion of patients, control was found to vary over the course of one day and, in some patients, from one minute to the next.

Our data strongly suggest that assessment of control at *a single point in time* does not accurately represent severity of IXT in many patients. Up to one half of patients showed marked changes in their level of control throughout the day. If a patient can be tropic in the morning and phoric in the afternoon, or vice versa, an isolated assessment of control should not be used to make clinical decisions such as whether or not to perform surgery. Our suggestion that assessment of control at a single point in time does not adequately represent severity is further supported by finding changes in control even within minutes in a small proportion of patients. A recent study by Romanchuk et al¹⁵ supports our assertion that a single measure of control may be unreliable. Analyzing consecutive clinic exams with descriptions of control, the authors report fluctuations from visit to visit and suggest that 'one cannot attach too much importance to measurements on any one visit'.

It may be that multiple measurements of control could be used to represent overall control in a particular patient; for example the sum or average of control scores assessed over the course of a day or a shorter period of time (for example, over an hour). Such a proposal would have to be evaluated in a multi-day study, to determine whether combined control scores (for example sum, mean, median) are reproducible from one day to another. Nevertheless, even if multiple measurements proved reproducible, bringing a patient back multiple times during a day to evaluate control would create practical challenges for clinical assessment.

Most other methods of rating control in IXT, whether semi-quantitative^{11, 13} or quantitative,⁹ use broader, less strictly-defined categories with perhaps less potential for observer variability. Nevertheless, our observation of a change from spontaneous tropia to phoria in several patients would have also constituted a 'change' in these alternative control scales. Petrunak et al¹² proposed a more finely graded control scale quantifying the time taken to refuse but it does not quantify any tropic phase so it cannot be applied to patients who have periods of tropia.

A compliment or alternative to clinic grading of control by an experienced observer is incorporating parental reports of the frequency with which the strabismus is manifest.⁴ The Newcastle Control Score⁹ includes a parental grade (0-3) of 'home control.' Other authors also recognize the role of parental estimates of control in the management of the condition.^{4, 5} Nevertheless, the reliability of such parental observations is unknown.

The findings of this study do not support the assertion that fatigue adversely affects control in patients with intermittent exotropia, a concept reported in previous studies.^{5, 16} While it may seem intuitive that fatigue would increase over the course of a day, none of our patients who showed a change in control over one day were worse at their last assessment. Interpretation of this observation is limited by the small number of patients (95% CI of n=6 is 0 to 46%). It is possible that control might only worsen when a patient is extremely tired and that this is rarely captured in a clinical assessment. Nevertheless, our data suggest that control is not universally worse at the end of the clinical day.

A change in control may be explained by factors other than fatigue. Parents will often describe noticing the exotropia when the child is not paying attention and Von Noorden has suggested that 'The extent to which an exodeviation is controlled by fusion depends not only on the size of the angle but also to a large extent on the general health, alertness, attention span and level of anxiety of the patient at the time of examination.'⁵ Increased sunlight, associated with monocular eye closure,¹⁷⁻¹⁹ is also thought to worsen control. We attempted to account for as many of these other factors as possible, by keeping the room illumination constant, standardizing the way the task was performed, and using a variety of interesting accommodative targets. When studying 'control', it is difficult to both standardize all measurements and still represent the clinical condition. By standardizing the measurements it is likely that the variability we observed is less than the actual variability of the condition.

Previous authors have suggested that poor control is a feature of large angle exodeviations.⁵ Our study does not support this assertion based on the median distance angle of deviation and median distance control score, measured through the course of a single day. Our previous study¹⁴ suggested that there was an overall positive correlation between angle and control score, but that study only documented a single measure of control. In both the present study, and in our previous study,¹⁴ there are notable examples of patients who have large angles and good control and smaller angles and poor control. For an individual patient, it is still not clear whether there is a relationship between angle of deviation and level of control. If a summary measure of control could be established, then this potential relationship would be a topic for further study.

Based on the results of our present intraday study, there may be 2 groups of patients with IXT differentiated by whether their control score did change during the day, which could be described as 'variable', or did not change during the day, which could be described as 'stable.' These potential sub-groups may have distinct clinical characteristics; for example, in our study the stable type appeared to have better distance control. Studying a larger cohort of patients would be needed to confirm the existence of 2 sub-groups, and conducting testing over a second day would clarify whether variability versus stability is a repeatable finding.

There are several limitations to this study. First, the sample size of patients observed over the course of a day was relatively small. Nevertheless, it is noteworthy that 6 (46%) of 13 patients showed a change of at least 2 levels in our control scale. The 95% confidence intervals of 6 of 13 are 19% and 75%, and a larger study would give a more robust estimate of the frequency of this finding. Second, our observation of the patient's control was of limited duration. A thirty second observation period followed by three 10-second periods of dissociation might not be expected to represent control over all waking hours. In addition, starting and stopping the clock at the beginning of the 30-second period will result in some variability based on the definition of the scale. The distinction between less than 50% tropic and greater than 50% tropic, and between greater than 50% tropic and constant tropia for 30 seconds, depends on "when you start the clock" and therefore changes between '3', '4' and '5' might be expected. Therefore, the definitions used in our control scale might erroneously lead to the impression of greater variability when the control was at the worse end of the spectrum. Nevertheless, we do not believe that this intrinsic variability on one part of our scale detracts from the primary conclusions of our study, that control varies throughout the day, since several subjects in our study changed from phoria to tropia and vice versa. The changes in the minute-to-minute component of this study may to some extent reflect the effect of dissociation, although in 1 of the 4 patients showing change, control was improved at the second assessment. Alternative measures of control are needed, involving prolonged observation or monitoring of the deviation, to provide a more representative summary of control in IXT.

The results of this study suggest that a single assessment of control cannot be relied upon to represent severity in an individual patient. Clinical decisions, such as whether or not to perform surgery, should not be made on the basis of a single observation of control. Apparent worsening or improvement of control, from one clinic visit to the next, may in fact represent normal variability of control in IXT. For control to be useful in defining severity of IXT, it seems highly likely that several assessments would be necessary to provide a representative summary measure. Additional studies are needed to determine the role of assessing control in the management of intermittent exotropia.

Acknowledgments

Supported by National Institutes of Health Grants EY015799 and EY011751 (JMH), Research to Prevent Blindness, Inc., 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.

References

1. Mohny BG, Huffaker RK. Common forms of childhood exotropia. *Ophthalmology* 2003;110:2093–6. [PubMed: 14597514]
2. Govindan M, Mohny BG, Diehl NN, Burke JP. Incidence and types of childhood exotropia. A population-based study. *Ophthalmology* 2005;112:104–8. [PubMed: 15629828]
3. Hatt S, Gnanaraj L. Interventions for intermittent exotropia. *Cochrane Database Syst Rev* 2006;19:CD003737. [PubMed: 16856017]
4. 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.
5. von Noorden, GK.; Campos, EC. Exodeviations. In: von Noorden, GK.; Campos, EC., editors. *Binocular Vision and Ocular Motility Theory and Management of Strabismus*. Vol. 6th. St. Louis: Mosby; 2002.
6. Wright, K. Exotropia. In: Wright, K., editor. *Pediatric Ophthalmology and Strabismus*. St Louis: Mosby Year Book; 1995.
7. Pratt-Johnson JA, Barlow JM, Tillson G. Early surgery in intermittent exotropia. *Am J Ophthalmol* 1977;84:689–94. [PubMed: 930997]

8. Abroms AD, Mohny BG, Rush DP, et al. Timely surgery in intermittent and constant exotropia for superior sensory outcome. *Am J Ophthalmol* 2001;131:111–6. [PubMed: 11162985]
9. Haggerty H, Richardson S, Hrisos S, et al. The Newcastle Control Score: a new method of grading the severity of intermittent distance exotropia. *Br J Ophthalmol* 2004;88:233–5. [PubMed: 14736781]
10. Mohny BG, Holmes JM. An office-based scale for assessing control in intermittent exotropia. *Strabismus* 2006;14:147–50. [PubMed: 16950743]
11. Chia A, Seenyen L, Long QB. A retrospective review of 287 consecutive children in Singapore presenting with intermittent exotropia. *J AAPOS* 2005;9:257–63. [PubMed: 15956946]
12. Petrunak, JL.; Rao, RC.; Baker, JD. The evaluation of office control in X(T): A systemic approach. In: de Faber, JT., editor. *Transactions of the 28th European Strabismological Association Meeting*. London: Taylor and Francis; 2004.
13. Stathacopoulos RA, Rosenbaum AL, Zanoni D, et al. Distance stereoacuity. Assessing control in intermittent exotropia. *Ophthalmology* 1993;100:495–500. [PubMed: 8479706]
14. Holmes JM, Birch EE, Leske DA, et al. New tests of distance stereoacuity and their role in evaluating intermittent exotropia. *Ophthalmology*. In Press
15. Romanchuk KG, Dotchin SA, Zurevinsky J. The natural history of surgically untreated intermittent exotropia--looking into the distant future. *J AAPOS* 2006;10:225–31. [PubMed: 16814175]
16. Burke MJ. Intermittent exotropia. *Int Ophthalmol Clin* 1985;25:53–68. [PubMed: 4077421]
17. Wang FM, Chryssanthou G. Monocular eye closure in intermittent exotropia. *Arch Ophthalmol* 1988;106:941–2. [PubMed: 3390058]
18. Costenbader F, Bair D, McPhail A. Vision in strabismus. *Arch Ophthalmol* 1948;40:438–53.
19. Campos EC, Cipolli C. Binocularity and photophobia in intermittent exotropia. *Percept Mot Skills* 1992;74:1168–70. [PubMed: 1501986]

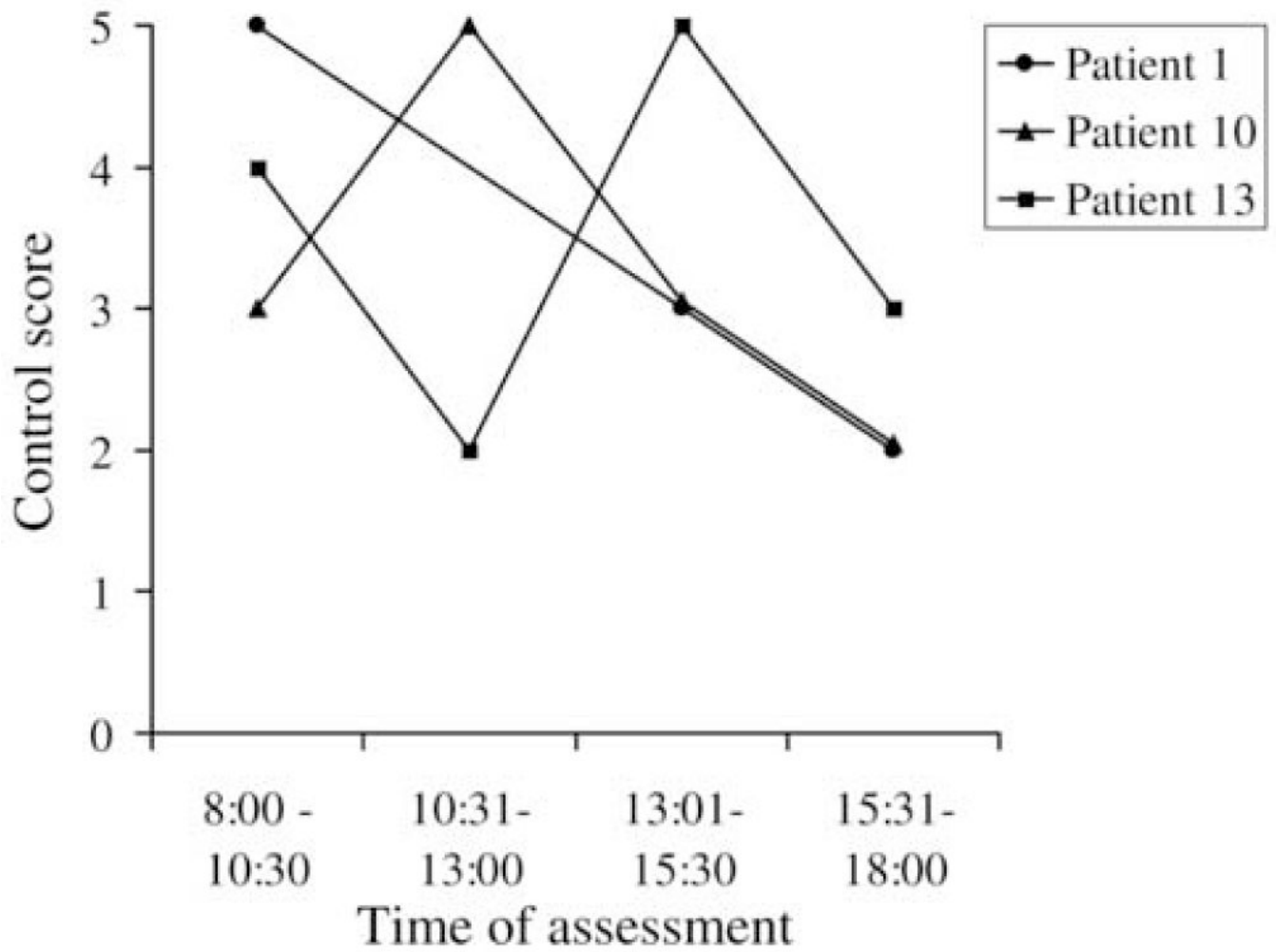


Figure 3.

Three patients showed change in distance control score over one day. All 3 patients showed changes between tropia (score 3-5) and phoria (score 0-2). No patient was worse at the end of the day.

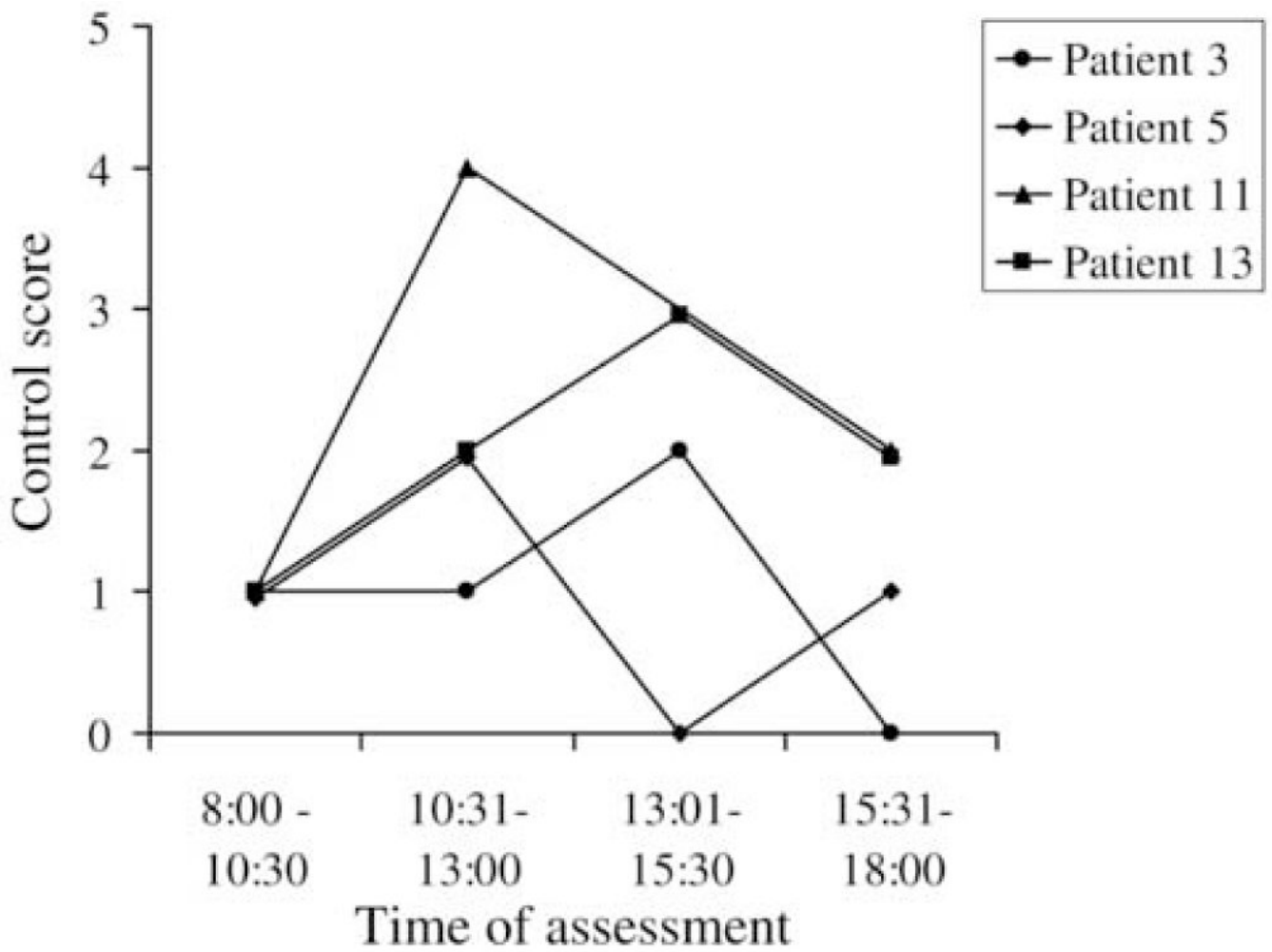


Figure 4. Four patients showed change in near control score over one day. 2 of these patients showed change from tropia (score 3-5) to phoria (score 0-2). No patient was worse at the end of the day.

Table 1

Clinic control scale description and instructions.*

<u>Control Score</u>	<u>Control Score Description</u>
5	Constant exotropia during a 30-second observation period (before dissociation)
4	Exotropia >50% of the time during a 30-second observation period (before dissociation)
3	Exotropia <50% of the time during a 30-second observation period (before dissociation)
2	No exotropia unless dissociated (10 seconds): recovery in > 5 seconds
1	No exotropia unless dissociated (10 seconds): recovery in 1-5 seconds
0	Pure phoria: < 1 second recovery after 10-second dissociation

Instructions:
The control scale is for both distance (3m) and near (1/3m) fixation. The fixation objects are accommodative and age appropriate, such as small stickers and videos for younger children and letters for older children and adults.
Levels 5 to 3 are assessed during a 30-second observation period. If exotropia is observed, testing stops and the control score is recorded as 5, 4, or 3 at that distance. If no exotropia is observed during the 30-second observation period, testing continues.
Levels 2 to 0 are then assessed and graded as the worst of three successive 10-second periods of occlusion. An occluder is placed over the *right* eye for 10 seconds and on removal the time required to re-fuse is noted. This process is repeated for the *left* eye and then a third occlusion trial is performed on the eye that required the longest time to re-fuse. The worst of these three 10-second trials is recorded, resulting in a control score of 2, 1, or 0 at that distance.
If the patient has a micro-esotropia by simultaneous cover test, but exodeviation by alternate cover test, the scale applies to the exodeviation.

* Mohney BG, Holmes JM. An office-based scale for assessing control in intermittent exotropia. *Strabismus* 2006;14:147-50.

Distance and near control scores as independently rated by 2 examiners at 2 time points within 5 minutes. Changes in control, between times one and two, of at least 2 levels are shown in bold type (available at <http://aaojournal.org>).

Table 2

Patient number	INVESTIGATOR 1		INVESTIGATOR 2		Near control		Time 1		Time 2	
	Distance	control	Distance	control	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
1*	4	2	1	1	4	2	1	1	1	1
2	2	1	1	1	2	1	1	1	1	1
3	2	2	2	2	2	2	2	2	2	2
4	2	2	1	1	2	2	2	1	1	1
5	2	2	1	1	2	2	2	1	1	1
6*	4	4	1	4	4	3	1	4		
7	1	1	0	1	1	1	1	1	1	1
8	5	5	1	1	5	5	1	1	1	1
9	2	2	0	0	2	2	0	0	0	0
10	5	5	1	1	5	5	1	1	1	1
11	1	2	0	1	1	2	0	1		
12*	3	3	1	3	3	3	0	3		
13*	4	5	2	4	4	5	2	4		
14	4	5	4	4	5	5	4	4		
15	2	1	2	1	1	1	1	2	1	1
16	4	4	4	4	4	4	4	4	4	4
17	2	1	1	1	2	1	1	1	1	1

* patients who showed a change in control

Table 3

Distance and near control scores over one day. Changes in control of at least 2 levels during the course of a day are shown in bold type (available at <http://aaojournal.org>).

Patient number	Distance control				Near control			
	Time 1 08:00-10:30	Time 2 10:31-13:00	Time 3 13:01-15:30	Time 4 15:31-18:00	Time 1 08:00-10:30	Time 2 10:31-13:00	Time 3 13:01-15:30	Time 4 15:31-18:00
1*	5	3	2	1	1	1	1	1
2	1	1	1	1	1	1	1	1
3*	2	3	2	3	1	1	2	0
4	2	2	2	2	1	1	1	1
5*	3	2	2	3	1	2	0	1
6	1	1	1	1	0	0	0	0
7	1	1	1	1	0	0	0	0
8	1	1	1	1	0	0	0	0
9	1	1	1	1	1	1	1	1
10*	3	5	3	2	1	1	1	1
11*	2	2	3	3	1	4	2	2
12	5	5	5	5	5	5	5	5
13*	4	2	5	3	1	2	3	2

* patients who showed a change in control