

Impossibility to eliminate observer effect in the assessment of adherence in clinical trials

Jonathan S Myers¹
Scott J Fudenberg¹
Robert E Fintelmann²
Lisa A Hark¹
Nitasha Khanna¹
Benjamin E Leiby³
Michael Waisbourd¹

¹Wills Eye Hospital, Glaucoma Research Center, Philadelphia, PA,
²Barnet Dulaney Perkins Eye Center, Phoenix, AZ, ³Division of Biostatistics, Thomas Jefferson University, Philadelphia, PA, USA

Purpose: To utilize the Travoprost Dosing Aid (DA) in the assessment of patient medication adherence, while also determining whether or not altering the functionality of the DA in three randomized subject groups can reduce observer effect.

Methods: Forty-five subjects were randomized into three groups: two with monitored DAs and one without monitoring. One group of subjects was given a DA that both monitored drop usage and had visual and audible alarms, while the other monitored group included subjects given a DA that had no alarms but continued to monitor drop usage. The third group was given a DA that had no alarm reminders or dose usage monitoring. Subjects were informed that some monitors would not be functional, in an attempt to reduce observer effect, or the effect of being monitored on subject behavior and adherence. A six-item questionnaire was also utilized to assess how the subjects felt about their adherence and DA use.

Results: The overall adherence rates were found to be 78% in the fully functional group (95% confidence interval: 70–88) and 76% in the no alarms group (95% confidence interval: 65–89). No association was seen between questionnaire response and medication adherence. The patients in the DA group without alarms had a significantly higher odds ratio of medication adherence if they reported on the questionnaire that using the DA did affect how much they used their drops.

Conclusion: Though the use of DA was expected to reveal different rates of adherence depending on the functionality of the DA between groups, patients with a nonfunctioning DA did not have a significant difference in medication adherence compared to those given a fully functional DA. This supports that an observer effect was not reduced despite these interventions, and that the subjects adhered to taking their medications as if they had a functioning DA and were being monitored.

Keywords: dosing aid, observer effect, glaucoma, adherence

Introduction

Glaucoma is a chronic, optic neuropathy characterized by optic nerve damage, visual field defects, elevated intraocular pressure, and progressive vision loss that impacts 60 million people worldwide.¹ Glaucoma management typically includes a daily eyedrop regimen.^{2,3} When properly used, eyedrop medication can effectively lower intraocular pressure, reduce optic nerve deterioration, preserve vision, and prevent glaucomatous blindness. However, rates of medication adherence and persistence (ie, continued use of medication over time) are especially low among patients diagnosed with glaucoma.^{4–6} Electronic monitoring of glaucoma medication administration found that adherence rates were poor.^{7–9} Okeke et al found that among patients being provided free medication for once-daily dosing who knew they were being monitored, 45% used their eyedrops less than 75% of the time. Further, close to one fifth of

Correspondence: Lisa A Hark
Wills Eye Hospital, 840 Walnut Street,
Suite 802, Philadelphia, PA 19107, USA
Tel +1 215 928 3045
Email lhark@willseye.org

Okeke et al's subjects (19.4%) used eyedrops less than 50% of the time.⁶ Prospective randomized trials on the impact of adherence on clinical outcomes in glaucoma are lacking.

Foucault wrote in his work on prison construction about the effect of monitoring on behavior.⁷ Ideally, he suggested, prisons should be constructed in such a way that monitoring is possible at all times but that the prisoner should not be able to tell at what point he is being monitored. He states that if a prisoner cannot tell whether or not he is being monitored, he will behave as if he is being monitored at all times. He coined the term "panopticism" for this effect.⁷

Clinical trials are performed to guide clinical practice, but the nature of a clinical trial may include biases that differ from clinical practice. Observer effect, reactivity, and "guinea pig" effect are some of the names given to the way observation influences the behavior of study subjects. Given that adherence to medications remains an important issue in medical treatment,^{2,3} it is important to explore the unobserved, or "real", medication adherence habits of patients with glaucoma. However, the evolution of ethical considerations and increasingly strict regulations governing clinical trials make it inappropriate to collect data on patients without their knowledge. With these considerations in mind, we designed a study to assess eyedrop medication adherence using a Travoprost Dosing Aid (DA), in an attempt to reduce observer effect.

Methods and design

Study organization

Forty-five subjects were randomly assigned to one of three groups by a predetermined three-way randomization chart: a "functional DA group" (Group 1) consisting of 20 participants who were given fully functioning DA devices with visual and audible alarms as dosing reminders and drop usage recording turned on (Alcon Laboratories, Inc., Fort Worth, TX, USA); a "no DA alarms group" (Group 2) consisting of 20 participants who were given DAs with disabled visual and audible alarms but that still monitored drop usage (ie, silent monitors); and a "nonfunctional DA group" (Group 3) consisting of five participants given nonfunctioning, placebo devices with visual and audible alarms disabled, and monitoring disabled. Subjects were informed that not all patients would be monitored. Groups 2 and 3 were given devices without audible or visual alarms to introduce doubt as to whether or not they were being monitored. All subjects were told at baseline that some of the devices were nonfunctional. Thus, the patients in Groups 2 and 3 were masked to whether or not they were being monitored on dose usage.

Patient study involvement

All subjects were given travoprost medication and trained on how to place the travoprost bottle in the DA and how to depress the lever arm to deliver a drop. Subjects were told that, when functioning, the device records usage when the lever is depressed. Subjects were asked to specify a 2-hour window during which they intended to use the medication each evening for which the functional DAs were set to trigger alarms. The subjects were supplied with free medication during the study period. Subjects returned in 6 weeks and the information was downloaded from each device and compiled in an identity-masked database. The subjects filled out a brief, non-validated six-item questionnaire on their perceptions of the device, including whether or not they believed their own device was functional. The questionnaire was intended to help us gain some potential insight into subjects' behaviors and perceptions after utilization of the DA.

Subjects and eligibility criteria

A total of 45 subjects were included, 20 in each monitored group and five in the unmonitored group. Inclusion criteria allowed for subjects with any type of glaucoma or glaucoma suspect diagnoses, treated with one or more glaucoma medications that included a topical prostaglandin analog. The number of patients chosen for each group was based on how many patients were interested in participating who fit the eligibility criteria for the duration of the study. The main exclusion criteria were mental and physical disabilities of subjects, including poor vision, precluding usage of the device and medication adherence. Patients were also excluded if they were unable to understand the study, if they did not instill their own drops, or if they were incapable of using the DA after a brief demonstration. All eligible patients had to be 18 years of age or older.

The study was reviewed by the Sidney Kimmel Medical College Institutional Review Board and deemed in concordance with the provisions of the Declaration of Helsinki and was registered in the NIH public database. Written informed consent was obtained.

DA

The DA can provide data only on use of the topical prostaglandin analog travoprost, because no other bottles for glaucoma medications fit within it. A bottle of travoprost is placed in the device and a lever is used to squeeze out a drop. A built-in memory chip records the time and date when the lever is depressed. The DA also has visual and audible reminders that can be set to remind patients to take their drops in a specified time period daily. Data can be downloaded to

assess whether or not a patient adhered to drop usage on a given day.¹ Because the device has the potential to make extra recordings when the lever is depressed accidentally, eg, if more than one dose is used to ensure instillation within the eye, more than one dose taken per eye per day was not counted in the adherence rate calculation (travoprost is indicated for once-daily use). When the lever was depressed outside the time window, it was assumed that a dose was not taken, and when the lever was depressed multiple times in the time window, only a single dose was assumed to have been delivered.¹⁰

Outcome measures

The primary outcome was medication adherence, defined as any use of eyedrops on a given day. The secondary outcome was awareness of monitoring. Medication adherence was based on the DA data and secondary questionnaire. This questionnaire was an instrument created for this study with no previous validity evidence, with adherence rated using a single yes/no question. The survey implemented patient awareness of monitoring as a second self-report measure that was documented by either a yes or no to each question.

Statistical analysis

Poisson regression with robust standard errors was used to model the relative risk of adherence while accounting for correlation among multiple measurements from the same subject. A first-order autoregressive structure was assumed for the working correlation structure.^{11,12} Logistic regression analysis was used to test for association between questionnaire response and adherence. The group to which subjects were assigned and their questionnaire responses were assessed and included as covariates in the logistic regression model to determine whether there was an association. Exact chi-square tests were used to look for differences among groups with respect to questionnaire responses. In a sensitivity analysis, both sets of models were repeated using only the first 15 days of follow-up to ensure equal amounts of data for each subject because some subjects did not complete the 6-week course. In this way, we also examined both adherence and short-term persistence. All analyses were performed using SAS version 9.1 (SAS Institute Inc., Cary, NC, USA).

Results

Baseline characteristics

Forty-five subjects (age: 67.6 years [standard deviation: 12.1]; 47% male) with a variety of glaucoma diagnoses

were recruited, more than half of which were on travoprost monotherapy prior to the study, the others were changed to travoprost for the study (Table 1). All but one (who died) completed the questionnaire. In the “functional DA group” one device malfunctioned. In the “no DA alarms group”, two devices were never returned despite persistent attempts to obtain these, and three devices malfunctioned. The malfunctioning devices showed no data recorded after the first day, and were returned to the manufacturer (Alcon Laboratories, Inc.) who was also unable to retrieve data from these or to determine the reason for the absence of data.

Patient adherence

Overall average medication adherence, defined as taking the eyedrops on a given day, was 78% of doses (95% confidence interval [CI]: 71%–85%). Adherence was nearly identical in both groups: 78% adherence in the “functional DA group” (95% CI: 70%–88%) and 76% adherence in the “no DA alarms group” (95% CI: 65–89). The relative risk of adherence comparing the “functional DA group” to “no DA alarms group” was 1.03 with a 95% CI of 0.85, 1.25 ($P=0.76$). In case reduced adherence over time in both groups was a factor, the first 15 days of therapy were assessed. Results differed slightly when considering only the first 15 days of data for each subject. The adherence for the first 15 days was 81% (95% CI: 74–89). Adherence did not significantly (statistically) differ between groups. There was 85% adherence in the “functional DA group” (95% CI: 78–93) and 76% adherence in the “no DA alarms group” (95% CI: 64–91). The relative risk of adherence comparing these two groups was 1.11 (95% CI of 0.91–1.36) ($P=0.29$). The functional

Table 1 Study group characteristics

| Characteristics | Functional DA group (Group 1) | No DA alarms group (Group 2) | Nonfunctional DA group (Group 3) |
|-----------------------------|-------------------------------|--|----------------------------------|
| Male/female patients | 8/12 | 9/11 | 4/1 |
| Age in years: mean \pm SD | 68.15 \pm 12.77 | 66.26 \pm 11.1 | 66.8 \pm 14.62 |
| Race: AA, C, A | 6, 13, 1 | 8, 11, 1 | 3, 1, 1 |
| POAG | 11 | 11 | 2 |
| Number of medications | | | |
| Only 1 | 9 | 17 | 1 |
| Two | 6 | 2 | 2 |
| Three | 5 | 1 | 2 |
| Dropout reasons | 1 died | 3 device malfunctions, 2 devices unable to recover data | none |

Abbreviations: SD, standard deviation; AA, African American; C, Caucasian; A, Asian; POAG, primary open angle glaucoma; DA, Travoprost Dosing Aid.

DA group had slightly but nonsignificantly higher adherence than the no DA alarms group.

Adherence within a 2-hour window of the time of planned dosing

Another planned evaluation was use of the medication within an hour before or after subject's chosen time of dosage. Overall adherence within this window was 42% (95% CI: 34–54). Adherence within 2 hours was higher in the fully functional group than in the no alarms group: 51% (95% CI: 40–65) vs 31% (95% CI: 20–48), although the difference was not statistically significant ($P>0.05$). The relative risk of adherence within 2 hours was 1.63 (95% CI: 1–2.68) ($P=0.052$). In the first 15 days, adherence within a 2-hour window of the planned dosage was 43% (95% CI: 33–56) for all groups together. Again, adherence was higher in the fully functional group (55%; 95% CI: 42–73) than in the alarms only group (29%; 95% CI: 18–47). The relative risk was 1.89 favoring greater adherence in the first group (95% CI: 1.08–3.30 [$P=0.025$]). These data suggest that the group with functional visible and audible alarms were more timely in their dosing in the first 15 days of the study.

Actual and self-reported medication adherence

Subjects in the “no DA alarms group” were more likely to admit to not administering eyedrops and were much less likely to agree that the DA affected how much they used their drops (see Table 2, questions 1 and 4). In Group 1, the “functional DA group”, 95% of subjects reported missing less than one drop per week on average, but the DA recordings showed only 30% to have missed less than one drop per week. For Group 2, the “no DA alarms group”, these numbers were 68% by self-report and 30% by DA.

Participants in the “nonfunctional DA group” were less likely to think that their DA was functioning (40%), although

the difference was not statistically significant ($P=0.20$), and most patients in the other two groups felt that they were being monitored (80% and 74%). Eighty percent of subjects in Group 3 felt the DA affected their drop use, which was statistically significant compared to the other two groups.

Patient adherence and the questionnaire responses

The increased adherence in patients who believed that they were being monitored was not statistically significant. There was no association between questionnaire response and adherence (see Table 3: odds ratios greater than 1 indicate that adherence was higher for those patients answering “yes” to the questions, while odds ratios less than 1 indicate that adherence was higher for patients answering “no” to the questions). The only statistically significant association was that patients in the “no DA alarms group” had a slightly higher odds ratio of adherence if they reported that using the DA affected how much they used their drops.

Discussion

It was initially hypothesized that adherence rates would be different in each group, depending on the modifications made to the DA and whether, as a result, a patient would feel as if he or she was being monitored. Adherence was not statistically different between the “functional DA group” with functioning visible and audible dosing alarm reminders, and the “no DA alarms group”. A similar percentage of patients suspected they were being monitored in both of these groups despite that those in Group 2 were given a DA that had no alarm reminders for the subjects. This supports that even without reminder alarms, subjects still adhered in a similar fashion to those with the reminder alarms. This, along with the questionnaire, substantiates that altering the presence of reminders on the DA did not significantly reduce subject perceptions of being observed/monitored.

Table 2 Association of group membership and questionnaire response

| Question | Number (%) with positive response | | | P-value for exact chi-square test of differences among three groups |
|--|-----------------------------------|----------------|---------------|---|
| | Group 1 (n=20) | Group 2 (n=19) | Group 3 (n=5) | |
| 1. Are you missing more than one drop per week on average? | 1 (5) | 6 (32) | 0 (0) | 0.056 |
| 2. Do you like the DA? | 17 (85) | 14 (74) | 5 (100) | 0.35 |
| 3. Do you feel as if you were monitored by the DA? | 16 (80) | 14 (74) | 2 (40) | 0.20 |
| 4. Did the DA affect your drop use? | 10 (50) | 4 (21) | 4 (80) | 0.029 |
| 5. Would you recommend the DA? | 17 (85) | 14 (74) | 5 (100) | 0.35 |
| 6. Would you continue to use the DA? | 16 (80) | 14 (74) | 5 (100) | 0.45 |

Notes: Group 1: fully functional; Group 2: no alarms; Group 3: nonfunctional.

Abbreviation: DA, Travoprost Dosing Aid.

Table 3 Association between responses to questionnaire and adherence by group

| Question | Group | Odds ratio (adherence vs no adherence – yes vs no) | 95% confidence interval | | P-value |
|---|-------|---|-------------------------|-------|---------|
| | | | Lower | Upper | |
| 1. Are you missing more than one drop per week on average | 1 | 1.11 | 0.62 | 1.97 | 0.73 |
| | 2 | 0.57 | 0.28 | 1.16 | 0.12 |
| 2. Do you like the DA? | 1 | 2.11 | 0.42 | 10.59 | 0.36 |
| | 2 | 1.95 | 0.39 | 9.79 | 0.42 |
| 3. Do you feel as if you were monitored by the DA? | 1 | 1.77 | 0.46 | 6.91 | 0.41 |
| | 2 | 4.97 | 0.71 | 34.78 | 0.11 |
| 4. Do you feel that the DA affected your drop use? | 1 | 1.33 | 0.45 | 3.88 | 0.60 |
| | 2 | 3.07 | 1.10 | 8.59 | 0.03 |
| 5. Would you recommend the DA? | 1 | 2.11 | 0.42 | 10.59 | 0.36 |
| | 2 | 1.95 | 0.39 | 9.79 | 0.42 |
| 6. Would you continue using the DA? | 1 | 2.11 | 0.42 | 10.59 | 0.36 |
| | 2 | 1.95 | 0.39 | 9.79 | 0.42 |

Notes: Group 1 (n=19): functional DA; Group 2 (n=15): no DA alarms.

Abbreviation: DA, Travoprost Dosing Aid.

According to the questionnaire, 74% of patients in Group 2 and 40% of patients in Group 3 felt they were being monitored (Table 2). Group 3 had a significant increase in medication adherence if they indicated that the DA improved their dose usage, though their DAs were nonfunctional. It appears that there was insufficient doubt about the monitoring process among subjects, despite their being informed at baseline that some subjects would not be monitored. Actions such as turning off the visual and audible DA alarms or providing nonfunctional devices did not affect subjects' perception of monitoring.

Lessons learned

This trial showed relationships between subject perception of monitoring and medication usage, an observer effect. Clinical trial results may be biased because subjects alter their behavior when they are monitored. In addition, patients' self-selection bias to enroll may also affect study results. These aspects should be considered when applying results to clinical practice. These issues have affected this study, despite efforts to convince subjects that they were not being monitored. This raises the possibility that a favorable outcome in a clinical trial may not directly translate into a favorable outcome for a patient in unobserved, "real" clinical practice.

Another significant finding is the difference in self-reported adherence by questionnaire compared to actual adherence as recorded by the DA. For instance, 95% of Group 1 subjects indicated that they had not missed more than one drop on average, per week, while the DA indicated that this actually only was true for 30% of these subjects. It would be interesting for additional studies to explore the difference in

self-reported adherence vs actual adherence and to assess in which populations those differences are the largest.

In previous studies, patients reported far higher medication use than their actual behavior. Several reasons have been suggested for this, including patients wanting to please their physicians, patients not wanting to admit an error, or patients not feeling comfortable enough to admit their concerns with the medication.⁹ Reported levels of non-adherence are affected by environmental cues and the method of questioning. Patient self-report and DA data produced different estimates of adherence in the current study and the self-report numbers for adherence were higher.

Although patient adherence can be assessed by indirect means (ie, interviewing, assessing pharmacy records), each of these has limitations. Electronic monitoring may be more accurate than any other option, but is also limited in that patients who know they are being monitored may change their behaviors as a result of the Hawthorne effect.^{13,14} Although subjects were aware for the entire study period that they may have been monitored, many clinical trials have actually found poor adherence despite patients' knowledge of monitoring, and often any effects that may be attributed to monitoring reactivity are transient. This could be explored in a study that follows patient adherence over a longer period of time than this study, to see if once subjects became accustomed to being monitored, they act naturally.^{13,14}

Interestingly, searches for other observation trials found behavioral studies of police that suggested that observation influences behavior but that over time this influence diminishes. This may be because the research subjects become accustomed to the observer and begin to act naturally. While

clearly a different situation, this may explain the drop-off in adherence over a longer period of time. The similarity between the groups independent of time may be due to the fact that all of them, as indicated in the questionnaire, felt monitored to some extent.^{13–15}

Limitations

The limitations of this study primarily were the small population sizes in each randomized group and that not all of the dosing aids were returned or able to provide the necessary information regarding dose usage. It was also a limitation to assess a 2-hour window of adherence and it would be interesting to see if a larger window correlated better with patient perception and adherence. Additionally, there was no true non-monitored comparison group in this study, because all participants believed it was possible that their daily medication use was monitored. It would be worth repeating the study under conditions where the DA is not also the method for collecting adherence data: if pharmacy fill data were available as the measure of adherence, some patients were also given a DA, and others were not. The confounding of the adherence measure with the monitoring device is a design problem that the current study could not solve. The self-report measure was a single, unvalidated item and more well-validated adherence self-report tools might produce different results. This study utilized the 2-hour window of patient adherence and questionnaire, but future studies may benefit by exploring additional metrics to assess adherence, such as other self-reported measures (such as the visual analog scale), or collecting the DA bottles and looking at the amount of medication left over.

Conclusion

It is both difficult and important to experiment with study designs that mitigate bias induced by artificial circumstances within a trial to achieve results that will reflect real clinical practice. This study suggests that the biases introduced by inclusion in a study may overwhelm deliberate attempts to induce doubt about observation.

Acknowledgment

The authors would like to acknowledge that this study was supported by a restricted grant from Alcon Laboratories, Inc. (Fort Worth, TX, USA).

Disclosure

There are no conflicts of interest from any of the contributing authors or affiliated institutions of this manuscript.

References

- Weinreb RN, Khaw PT. Primary open-angle glaucoma. *Lancet*. 2004;363(9422):1711–1720.
- Blackwell B. Treatment adherence. *Br J Psychiatry*. 1976;129:513–531.
- Osterberg L, Blaschke T. Adherence to medication. *N Engl J Med*. 2005;353(5):487–497.
- Friedman DS, Okeke CO, Jampel HD, et al. Risk factors for poor adherence to eyedrops in electronically monitored patients with glaucoma. *Ophthalmology*. 2009;116(6):1097–1105.
- Wagner GJ, Ghosh-Dastidar B. Electronic monitoring: adherence assessment or intervention? *HIV Clin Trials*. 2002;3(1):45–51.
- Okeke CO, Quigley HA, Jampel HD, et al. Adherence with topical glaucoma medication monitored electronically the Travatan Dosing Aid study. *Ophthalmology*. 2009;116(2):191–199.
- Foucault M. *Discipline & Punish: The Birth of the Prison*. New York: Vintage Books; 1995.
- Kass MA, Meltzer DW, Gordon M. A miniature compliance monitor for eyedrop medication. *Arch Ophthalmol*. 1984;102(10):1550–1554.
- Kass MA, Meltzer DW, Gordon M, Cooper D, Goldberg J. Compliance with topical pilocarpine treatment. *Am J Ophthalmol*. 1986;101(5):515–523.
- Friedman D, Jampel H, Congdon NG, Miller R, Quigley HA. The TRAVATAN Dosing Aid accurately records when drops are taken. *Am J Ophthalmol*. 2007;143(4):699–701.
- Robin AL, Novack GD, Covert DW, Crockett RS, Marcic TS. Adherence in glaucoma: objective measurements of once-daily and adjunctive medication use. *Am J Ophthalmol*. 2007;144(4):533–540.
- Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol*. 2004;159(7):702–706.
- Gittelsohn J, Shankar AV, West KP Jr, Ram R, Gnywali T. Estimating Reactivity in Direct Observation Studies of Health Behaviors. *Hum Organ*. 1997;56(2):182–189.
- Kalichman SC, Amara CM, Swetzes C, et al. A simple single-item rating scale to measure medication adherence: further evidence for convergent validity. *J Int Assoc Physicians AIDS Care (Chic)*. 2009;8(6):367–374.
- Kass MA, Gordon M, Morley RE Jr, Meltzer DW, Goldberg JJ. Compliance with topical timolol treatment. *Am J Ophthalmol*. 1987;103(2):188–193.

Patient Preference and Adherence

Publish your work in this journal

Patient Preference and Adherence is an international, peer-reviewed, open access journal that focuses on the growing importance of patient preference and adherence throughout the therapeutic continuum. Patient satisfaction, acceptability, quality of life, compliance, persistence and their role in developing new therapeutic modalities and compounds to optimize

Submit your manuscript here: <http://www.dovepress.com/patient-preference-and-adherence-journal>

Dovepress

clinical outcomes for existing disease states are major areas of interest for the journal. This journal has been accepted for indexing on PubMed Central. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.