

Residency Training: The King-Devick test and sleep deprivation

Study in pre- and post-call neurology residents

Emma C. Davies, BS
Sam Henderson
Laura J. Balcer, MD,
MSCE
Steven L. Galetta, MD

Correspondence & reprint
requests to Dr. Balcer:
lbalcer@mail.med.upenn.edu

ABSTRACT

Objective: The current study investigates the effect of sleep deprivation on the speed and accuracy of eye movements as measured by the King-Devick (K-D) test, a <1-minute test that involves rapid number naming.

Methods: In this cohort study, neurology residents and staff from the University of Pennsylvania Health System underwent baseline followed by postcall K-D testing ($n = 25$); those not taking call ($n = 10$) also completed baseline and follow-up K-D testing. Differences in the times and errors between baseline and follow-up K-D scores were compared between the 2 groups.

Results: Residents taking call had less improvement from baseline K-D times when compared to participants not taking call ($p < 0.0001$, Wilcoxon rank sum test). For both groups, the change in K-D time from baseline was correlated to amount of sleep obtained ($r_s = -0.50$, $p = 0.002$) and subjective evaluation of level of alertness ($r_s = 0.33$, $p = 0.05$) but had no correlation to time since last caffeine consumption ($r_s = -0.13$, $p = 0.52$). For those residents on their actual call night, the duration of sleep obtained did not correlate with change in K-D scores from baseline ($r_s = 0.13$, $p = 0.54$).

Conclusions: The K-D test is sensitive to the effects of sleep deprivation on cognitive functioning, including rapid eye movements, concentration, and language function. As with other measures of sleep deprivation, K-D performance demonstrated significant interindividual variability in vulnerability to sleep deprivation. Severe fatigue appears to reduce the degree of improvement typically observed in K-D testing. **Neurology® 2012;78:e103-e106**

GLOSSARY

K-D = King-Devick; **KSS** = Karolinska Sleepiness Scale.

Sleep deprivation has been demonstrated to negatively impact multiple aspects of neurocognition, including diminished attention, altered perception, impaired memory, and slowed visuomotor response.¹ Despite evidence that sleep deprivation significantly hampers neurocognitive skills, the degree of potential compensation for cognitive slowing continues to be debated. In a study of surgical residents, no abnormalities in arithmetic calculations and surgical knot tying were observed after overnight sleep deprivation.² Furthermore, substantial interindividual differences in vulnerability to sleep loss effects have been demonstrated, as a study of volunteers undergoing sleep deprivation periods ranging from 1 to 3 days showed some participants had minimal change in cognitive functioning while others had significant impairment on a battery of neurocognitive tests.³

Prior research has demonstrated that fatigue caused by sleep deprivation has a profound impact on eye movements, particularly by slowing peak saccadic velocity,⁴ increasing spontaneous blinking,⁵ and diminishing accuracy of smooth pursuit.⁶ The King-Devick (K-D) test involves rapid number naming and captures impairment of eye movements, attention, language, and other correlates of suboptimal brain function. The K-D test has recently been verified as a method to screen for

From the Departments of Neurology (E.C.D., S.H., L.J.B., S.L.G.), Ophthalmology (L.J.B., S.L.G.), and Epidemiology (L.J.B.), University of Pennsylvania School of Medicine, Philadelphia.

Study funding: Supported by NIH K24 EY 018136 (Dr. Balcer).

Go to Neurology.org for full disclosures. Disclosures deemed relevant by the authors, if any, are provided at the end of this article.

concussion in athletes, with increased times associated with cognitive slowing and worse concussion outcomes.^{7,8}

The current research investigates the impact of sleep deprivation on eye movements and attention as measured by the K-D test in residents taking overnight call. It aims to study some unresolved questions surrounding the effects of sleep deprivation, including better identification of interindividual differences in vulnerability to sleep loss effects and quantification of the role of fatigue on K-D testing. The study hypothesizes that increased sleep deprivation will be associated with increased time and increased errors (worse performance) on the K-D test.

METHODS Study participants. Neurology residents and staff from the University of Pennsylvania Health System were enrolled in this study. Subjects are a convenience sample of individuals who were taking overnight call ($n = 25$) and individuals who were not taking call ($n = 10$) who were willing to undergo testing.

Standard protocol approvals, registrations, and patient consents. Study protocols were approved by the University of Pennsylvania Institutional Review Board. All participants signed written informed consent.

Participant survey. Before testing, each participant completed a survey to self-report the number of hours of sleep obtained in the prior 24 hours, approximate number of hours of sleep over the past week, timing of caffeine consumption, and evaluation of sleepiness. The Karolinska Sleepiness Scale (KSS), a commonly utilized assessment of overall tiredness based on a 9-point scale (1 = very alert, 3 = alert, 5 = neither alert nor sleepy, 7 = sleepy [but not fighting sleep], 9 = very sleepy [fighting sleep]), was used to measure the level of sleepiness in participants.⁹ Also, reports of the prevalence of sleep deprivation symptoms, including falling asleep during meetings and problems with memory and concentration, were obtained.

The K-D test. The K-D test is based on the time to perform rapid number naming.⁷ It involves reading aloud a series of single-digit numbers from left to right on 3 test cards. All participants were given the same standardized instructions prior to each testing session, specifically to read the numbers as quickly and accurately as possible. The K-D test yields scores for time (sum of number of seconds required to read each test card) and errors (total number of mistakes made on the test cards).

Each participant completed 2 testing sessions, 1 as a baseline and 1 as follow-up, typically within a 24- to 72-hour time period. The first session was performed in residents and staff members when they were not postcall and after ≥ 6 hours of sleep. In residents taking call, the second session was completed the morning after overnight call. In residents and staff members not taking overnight call, the second session was completed at the same time of day after ≥ 6 hours of sleep. All testing was performed by a medical student blinded to the results of the first session.

Statistical analysis. Data analyses were performed using Stata 12.0 software (StataCorp, College Station, TX). Differences in follow-up and baseline times were compared for participants currently taking call and those not taking call using the Wilcoxon rank sum test. Associations between change in K-D test performance and duration of sleep during the call period, caffeine consumption, and subjective scoring of sleepiness were evaluated using Spearman rank correlations. For all statistical tests, type I error for significance was set at $p < 0.05$.

RESULTS Characteristics and K-D testing data for the resident and staff cohort are summarized in the table. The participants taking call did not have less sleep over the prior 24 hours at baseline ($p = 0.82$) but did have less sleep when postcall relative to the control group ($p < 0.0001$, Wilcoxon rank sum test).

Less improvement from baseline K-D times were observed in residents taking call when compared to residents and staff not taking call ($p < 0.0001$, Wilcoxon rank sum test; figure, A). Errors on K-D testing were minimal throughout the study (total of 7 errors), but were more frequent among residents taking call both at baseline (1 of 25 participants with at least 1 error) and at follow-up (4 of 25 with at least 1 error).

Changes in K-D time scores from baseline were correlated to amount of sleep obtained (for the entire cohort including subjects who were postcall and those not taking call), with less sleep associated with less improvement in K-D score ($r_s = -0.50$, $p = 0.002$). Within the postcall resident group, however, the magnitude of changes in K-D scores from baseline did not correlate with duration of sleep on call, although performance was impaired in this group. Changes in K-D time score from baseline were also modestly correlated to the subjective evaluation of sleepiness at time of testing, with greater improvement from baseline among those who reported less sleepiness ($r_s = 0.33$, $p = 0.05$; figure, B). There was no relation between changes in K-D time scores from baseline and time of last caffeine consumption ($r_s = -0.13$, $p = 0.52$) or length of sleep while on call among postcall residents ($r_s = 0.13$, $p = 0.54$).

DISCUSSION The K-D test is sensitive to the impact of sleep deprivation on cognitive functioning, including aspects such as rapid eye movements, concentration, and language. Participants not taking call showed a median improvement of 3.8 seconds in follow-up K-D scores, consistent with the learning effect previously described in K-D testing.^{7,8} However, residents had a median slowing of about 0.23 seconds on postcall K-D testing, suggesting that the learning effect was negated by sleep deprivation. While this worsening in K-D test performance is not

Table Characteristics and scores for the neurology resident and staff cohort

	Residents taking call (n = 25)	Residents and staff not taking call (n = 10)
Age, y, mean \pm SD	30 \pm 3	31 \pm 10
Sex, n (% male)	9 (36)	5 (50)
Duration of sleep in past 24 hours at baseline, h, median (range)	7.5 (6.0 to 12.0)	7.0 (6.0 to 13.0)
Duration of sleep in past 24 hours at time of follow-up, h, median (range) ^a	2.0 (0 to 6.5)	6.75 (6.0 to 8.0)
KSS at baseline, median (range) ^b	3 (1 to 7)	4 (3 to 5)
KSS at follow-up, median (range) ^b	5 (1 to 9)	3 (1 to 7)
Time since caffeine intake at baseline, h, mean \pm SD	2.2 \pm 1.7	2.3 \pm 1.9
No. of times subject has fallen asleep during a meeting/activity, median (range) ^a	1.65 (0 to 5)	0.7 (0 to 5)
Baseline K-D time score, s, median (range)	36.9 (26.0 to 46.9)	38.6 (27.3 to 43.7)
Follow-up K-D time score, s, median (range)	37.4 (25.6 to 49.7)	34.9 (25.5 to 38.1)
Change in K-D time score from baseline, s, median (range) ^{a,c}	0.23 (-4.6 to 8.6)	-3.8 (-7.7 to -1.8)
Participants with at least 1 error on K-D testing at baseline, n (%)	1 (4%)	0 (0%)
Participants with at least 1 error on K-D testing at follow-up, n (%)	4 (16%)	0 (0%)

Abbreviations: K-D = King-Devick test; KSS = Karolinska Sleepiness Scale.

^a Indicates a variable that was statistically significantly different between the 2 groups: $p < 0.0001$ for duration of sleep at time of follow-up; $p = 0.02$ for number of times fallen asleep; $p < 0.0001$ for change in K-D score from baseline.

^b KSS is a subjective measure of overall tiredness with rating scale 1–9; higher scores represent greater levels of sleepiness.

^c Change in K-D time score from baseline calculated as (follow-up time) – (baseline time); negative numbers indicate improvement in score, consistent with mild learning effect.

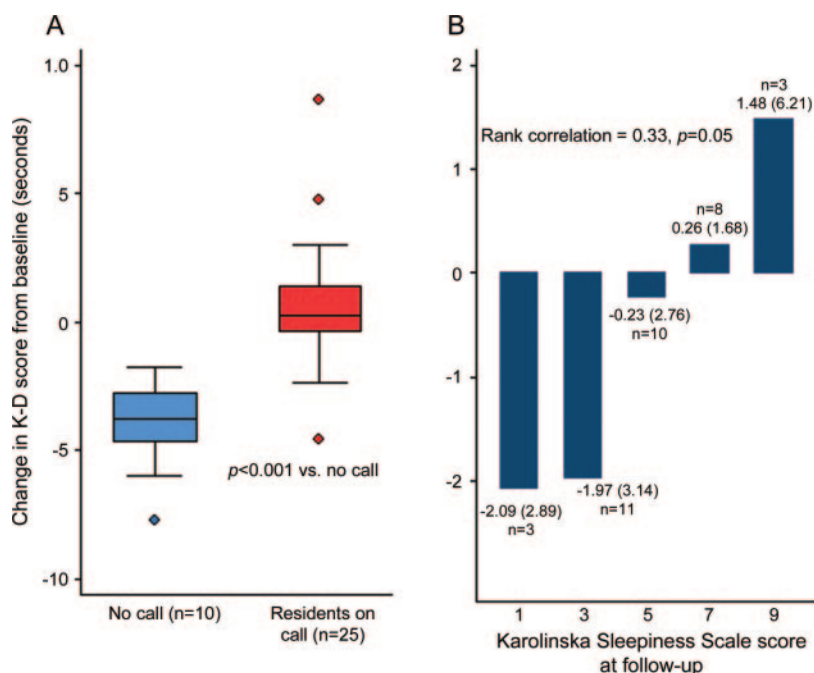
as extreme as seen in a cohort of athletes with concussions (reported median 5.9-second increase in K-D time),⁸ it demonstrates the potential effectiveness of K-D testing to detect eye movement slowing in sleep deprivation. Furthermore, postcall residents had an increased number of errors on follow-up relative to control subjects, similar to the increased number of errors observed in athletes with concussions.⁸ These results suggest that increased time and error on K-D testing capture the deleterious effects of sleep deprivation on eye tracking and attention.

Less sleep in the prior 24 hours was associated, in the overall cohort of on-call and noncall participants, with less improvement in K-D time from baseline. However, among residents taking call, no association was found between duration of sleep obtained on-call and K-D follow-up performance. Nonetheless, the performance on the K-D test was impaired in this group. In other words, postcall residents obtained less sleep time than subjects not taking call, which correlated to poor performance on the K-D test, but when a subanalysis of the postcall residents was completed no correlation between K-D time from baseline and sleep was observed. The finding is unlikely to be due to variability in test performance; in fact, recent studies demonstrated a high degree of test-retest reliability for K-D with intraclass correlations of 0.97.⁷ The result could be explained, however, by

interindividual vulnerability to the effects of sleep deprivation that has been documented with other sleep deprivation measures.^{1,3} In the current study, some residents obtaining zero hours of sleep on call improved at a level similar to controls while others were significantly impaired in their K-D performance even when obtaining several hours of sleep on call. No study variable (including age, sex, time to caffeine, level of training, hospital service, or prevalence of sleep deprivation symptoms) effectively predicted which individual would be more affected by sleep deprivation. One potential confounding factor may have been the amount of caffeine consumption prior to testing; while our study captured the time since last caffeine intake, the actual quantities of caffeine were not evaluated in this study.

Self-report of increased level of sleepiness (using the KSS) was associated with less improvement in K-D times from baseline. This finding corresponds to prior research that found participants reporting an increased level of sleepiness had slower peak eye saccade velocities relative to participants rating themselves as more alert.⁵ Although the ability to appropriately identify level of alertness has been debated, recent evidence suggests that subjective alertness and performance are modestly correlated, with most discrepancy during the biological night.¹⁰ The association between subjective level of sleepiness and

Figure Change in King-Devick (K-D) test time from baseline without call vs postcall (A) and mean change in K-D test time from baseline vs Karolinska Sleepiness Scale (KSS) score (B)



(A) Change in K-D test time from baseline without call vs postcall. Change in K-D test time from baseline in participants not taking call vs residents taking call. Box plots show the change in K-D time from baseline (seconds) in residents and staff not taking call ($n = 10$) compared to residents taking call ($n = 25$). The lines in the boxes represent the medians, and boxes delineate the interquartile range (25th to 75th percentiles). Whiskers represent the range of observations minus outliers; the circles represent outliers. Change in K-D time calculated by (follow-up time) – (baseline time); negative numbers indicate improvements in scores, consistent with mild learning effect. Participants taking call had less improvement in K-D scores relative to those not taking call ($p < 0.0001$; Wilcoxon rank sum test). (B) Mean change in K-D test time from baseline vs subjective evaluation of sleepiness with KSS. The KSS is a Likert scale with ratings from 1 to 9, with higher numbers representing greater levels of sleepiness. Rank correlations of change in K-D score from baseline vs KSS score at follow-up were modest but in the direction of greater improvement with less sleepiness ($r_s = 0.33$, $p = 0.05$ for the entire cohort [$n = 35$]).

improvement in K-D test times in the current study provides further support for the concept that individuals are fairly effective in rating their level of alertness. However, self-report of sleepiness is a subjective measure not capable of capturing individual vulnerability to sleep deprivation effects.

Severe fatigue (ratings sleepy and very sleepy on the KSS) significantly affected follow-up performance on the K-D test in the present study. Since prior research has demonstrated no effect on K-D performance from athletic workout fatigue,⁸ the current results extend our understanding of K-D testing in relation to extreme fatigue levels. In the search to objectively evaluate the effects of sleep deprivation, the K-D test offers a simple and quick method to measure degree of eye movement slowing in subjects. Further research with larger cohorts is needed to ex-

pand the subject sample size, elaborate interindividual variability in vulnerability to the effects of sleep deprivation, and evaluate the impact of different sleep loss patterns on attention and eye movements. Indeed, the K-D test offers the potential to monitor resident performance under a variety of call schedules, including night float systems, and to test the association between eye movement slowing and clinical errors.

AUTHOR CONTRIBUTIONS

E.C. Davies: collection of data, analysis and interpretation of data, drafting and revising the manuscript. S. Henderson: collection of data, revising the manuscript. Dr. Balcer: design and conceptualization of study, analysis and interpretation of data, drafting and revising the manuscript. Dr. Galetta: design and conceptualization of study, interpretation of data, revising manuscript.

DISCLOSURE

E.C. Davies and S. Henderson report no disclosures. Dr. Balcer has received honoraria and consulting fees from Biogen Idec, Novartis, and Vaccinex. Dr. Galetta has received honoraria for speaking from Biogen Idec, Novartis, and Teva. [Go to Neurology.org for full disclosures.](#)

REFERENCES

1. Durmer JS, Dinges DF. Neurocognitive consequences of sleep deprivation. *Semin Neurol* 2005;25:117–129.
2. Leff DR, Orihuela-Espina F, Athanasiou T, et al. Circadian cortical compensation: a longitudinal study of brain function during technical and cognitive skills in acutely sleep-deprived surgical residents. *Ann Surg* 2010;252:1082–1090.
3. Van Dongen HPA, Baynard MD, Maislin G, Dinges DF. Systematic inter-individual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep* 2004;27:423–433.
4. Fransson PA, Patel M, Magnusson M, Berg S, Almladh P, Gomez S. Effects of 24-hour and 36-hour sleep deprivation on smooth pursuit and saccadic eye movements. *J Vestib Res* 2008;18:209–222.
5. Goldich Y, Barkana Y, Pras E, Zadok D, Hartstein M, Morad Y. The effects of sleep deprivation on oculomotor responses. *Curr Eye Res* 2010;35:1135–1141.
6. Zils E, Sprenger A, Heide W, Born J, Gais S. Differential effects of sleep deprivation on saccadic eye movements. *Sleep* 2005;28:1109–1115.
7. Galetta KM, Barrett J, Allen M, et al. The King-Devick test as a determinant of head trauma and concussion in boxers and MMA fighters. *Neurology* 2011;76:1456–1462.
8. Galetta KM, Brandes LE, Maki K, et al. The King-Devick test and sports-related concussion: study of a rapid visual screening tool in a collegiate cohort. *J Neurol Sci* 2011;309:34–39.
9. Akerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. *Int J Neurosci* 1990;52:29–37.
10. Zhou X, Ferguson SA, Matthews RW, et al. Mismatch between subjective alertness and objective performance under sleep restriction is greatest during the biological night. *J Sleep Res Epub* 2011 May 13.