Sleep, 16:S25–S29

Natural History of Sleep Disordered Breathing in Community Dwelling Elderly

*Sonia Ancoli-Israel, *Daniel F. Kripke, †Melville R. Klauber, *Linda Parker, *Carl Stepnowsky, *Allison Kullen and *Robert Fell

**Department of Psychiatry, University of California, San Diego, Veterans Affairs Medical Center, La Jolla, California, U.S.A.*
†Department of Community and Family Medicine, University of California, San Diego, California, U.S.A.

Summary: Mild sleep disordered breathing is very common in the elderly, but little is known about the course of the disorder over time. Twenty-four elderly people from a population-based study were recorded three times over an 8.5-year period. There were no significant changes in either apnea index or in respiratory disturbance index (RDI) over time, even when controlled for body mass index. For most subjects, there was great variability over time in the number of respiratory disturbances. The sensitivity of RDI ≥ 15 at visit 1 predicting RDI ≥ 15 at visit 3 was only 20%. The predictive value was 50%. Sleep disordered breathing measured at a single point in time is rather weakly predictive of the severity of breathing disorder 4–8 years later. **Key Words:** Sleep disordered breathing—Elderly—Mortality—Natural history—Longitudinal studies.

Very little is known about the natural history of sleep disordered breathing (SDB). A few studies have looked at night-to-night variability of SDB. Wittig et al. (1) found a reliability of 0.86 for apnea indices among patients with >100 apneas a night, but a much lower correlation of 0.33 among patients with <100 apneas per night. Mosko et al. (2) found greater night-to-night variability with only 43% of these patients meeting criteria for sleep apnea on night one, also meeting criteria again on following nights. Aber et al. (3) found no change in the mean apnea index from night to night, but did find that 35% would have changed classification. Others have also reported night-to-night variability of sleep apnea (4–6). Ancoli-Israel and colleagues found that correlations of sleep apnea from one night to another were $r_s = 0.80$ for hospital in-patients and $r_s = 0.73$ for nursing home patients. Lee and Giblin, on the other hand, reported that in healthy men, one night was sufficient to rule out sleep apnea (7). Given the night-to-night variability of sleep apnea, it becomes even more important to understand what happens to the severity of the disorder over longer time periods.

Few studies have reported data on the natural history of sleep apnea. Correlations over time would reflect both the day-to-day and year-to-year variance. Rosenthal et al. (8) followed 23 patients with mild sleep apnea for 2 years and found that they showed improvement in their apnea index, independent of weight change. Rosenthal et al. (8) concluded that in the absence of weight gain, mild SDB is not a progressive disease.

Bliwise et al. (9) followed elderly subjects for 2.8 years and middle-aged adults for 8.1 years. There was a statistically significant increase of 2.3–3.8 apnea points in both groups, but such a small increase would not be clinically significant. Mason et al. (10) reanalyzed the Bliwise data and found the correlation of sleep apnea over time was only 0.37 (not significant) for the middle-aged group and 0.64 ($p < 0.01$) for the group of elderly. Phoha et al. (11) followed older community residents with mild sleep apnea at initial evaluation and found a statistically significant, but clinically mild increase in respiratory disturbance index over 3 years. Bliwise et al. (12) found similar results in older individuals followed over 5 years. Zorick et al. (13), on the other hand, followed patients with mild sleep apnea over 2 years and found that, independent of weight, apnea indices were significantly reduced, indicating an improvement in the syndrome. Ancoli-Israel et al. (14) followed elderly patients over 2–3 years and found that 20% showed an increase in apnea index, 20% showed a decrease and 60% remained the same. Some of the differences among samples reflect regression to the mean among samples with high initial indices.

Mason et al. (10) followed a group of 30 elderly subjects, also with mild SDB. These subjects were originally part of a larger study of prevalence of sleep apnea in the elderly (15), which began in 1981. Subjects were randomly selected, and sleep was recorded in the subjects' own homes with unattended monitoring. Four and a half years later, the 30 subjects were reevaluated with the same technology (10). There was no significant difference (either increase or decrease) in either apnea index or in respiratory disturbance index (RDI). The correlation between the RDI at initial evaluation and 4.6 years later was 0.69 ($p < 0.001$). The correlation for apnea index at initial evaluation and 4.6 years later was 0.50 ($p < 0.05$). Thirty-six percent of the original 30 had met criteria for sleep apnea at the initial evaluation. Only 16% of the original 30 elderly met criteria for sleep apnea at follow-up. Such data suggested that spontaneous remission may have been occurring. We now have the opportunity to report on the third recording of this same group of subjects.

METHODS

Subjects

Subjects were originally 427 men and women 65 years old and over, who had been randomly selected and interviewed by telephone between 1981 and 1985. Each had been interviewed about sleep and health and was recorded for one night. The original sample had a mean age of 72.5 years when first recorded (SD = 6.1; range = 65–95) (15). Approximately 4.6 years later, 134 were recontacted and 40 (30%) agreed to be restudied. All 40 subjects who completed a second recording were invited to be restudied a third time. The protocol was approved by the Committee on Investigations of Human Subjects of the University of California, San Diego.

Apparatus

Each elderly volunteer was studied with the modified Respitrace/Medilog portable recording system (16), which recorded two channels of respiration (thoracic and abdominal), one channel of tibialis electromyograph (EMG) and one channel of wrist activity (to distinguish wake from sleep). Data were stored on an analog tape-recorder and played back onto a Grass model 78 polygraph. This methodology, previously validated against traditional polysomnography, shows high reliability: apnea index, $r_s = 0.80$ ($p < 0.01$), myoclonus index, $r_s = 0.64$ ($p < 0.005$), total sleep time, $r_s = 0.82$ ($p < 0.01$), and wake after sleep onset, $r_s = 0.61$ ($p < 0.01$) (17).

On this third follow-up visit, blood oxygen saturation levels were also recorded for one of the two nights of recording. A portable finger pulse oximeter (Ohmeda 3700) and portable notebook computer were used. The oximeter sampled blood oxygen saturation levels every 2 seconds throughout the night and, to minimize loss of data, stored the data in memory and on disk. PROFOX, an oximetry program, was used to compute the number of desaturation events >4% and lasting <3 minutes, the mean durations of desaturation events, and the percent time spent at different saturation levels (18).

Procedure

Visit 1 occurred between 1981 and 1985. Visit 2 occurred between 1985 and 1988. Visit 3 occurred in 1990 or 1991. Each of the 40 original subjects seen at visit 2 was recontacted for visit 3. Twenty-four were available for re-study. Once they agreed to continued participation, they were re-interviewed about sleep and medical history, including medication use and any change in medical status since the previous visit. The Geriatric Depression Scale (19,20) and the Mini-Mental Status Examination (21) (which were not previously used) were also administered at visit 3.

Subjects were recorded in their own homes with the modified Respitrace/Medilog system. Twenty subjects were recorded for two nights but four could only be recorded for one night. Oximetry, which was not available at visits 1 or 2, was recorded during visit 3 on one of the two nights. Each sleep record was scored for total sleep time (TST), wake time (WASO), number of awakenings, number and type of apneas, number and type of hypopneas and number of leg jerks. Apnea index (AI—number of apneas per hour of sleep), respiratory disturbance index (RDI—number of apneas and hypopneas, or number of respiratory disturbances per hour of sleep) and myoclonus index (number of leg jerks per hour of sleep) were computed. When two nights of data were available, the mean of the two nights was used.

Data analyses

Changes in AI, RDI and other variables over time were analyzed using the restricted maximum likelihood (REML) method (22). This method allows random rates of change, unequally spaced observations and missing data. Pearson correlations were used to examine the consistency of variables over time.

RESULTS

Compliance

Of the 40 subjects who completed a second follow-up recording, 24 (60%), (10 women and 14 men) agreed to be studied a third time, 11 (28%) refused further

TABLE 1. *Respiratory indices by visit*

	Visit 1 (n = 24)	Visit 2 (n = 23)	Visit 3 (n = 23)
AI mean	3.27	2.76	3.07
SD	3.22	4.59	7.17
Median	2.24	1.06	0.90
Range	0–10.55	0–17.34	0–35.00
RDI mean	13.41	13.41	10.44
SD	9.16	9.78	9.27
Median	11.30	10.86	7.34
Range	2.06–38.25	0–38.58	0.68–42.07

Abbreviations: AI = apnea index, SD = standard deviation, RDI = respiratory disturbance index.

participation and five (12%) had died. The 11 subjects who refused a third study did not differ from the subjects who completed visit 3 in mean apnea index, in mean respiratory disturbance index or in general health status. Due to equipment failure, data on one subject was lost at visit 2 and data on another subject was lost for visit 3.

Demographics

Visit 3 occurred a mean of 3.6 years (SD = 1 year, range = 2.1–5.2 years) after visit 2 and 8.5 years (SD = 0.4 year, range = 7.6–9.4 years) after visit 1. At the time of visit 1, the mean age of the 24 subjects was 70.1 years (SD = 3.0, range = 65–77). At the time of visit 2, the mean age of the 24 subjects was 75.0 years (SD = 2.9, range = 71–83). At the time of visit 3, the mean age of the subjects was 78.7 years (SD = 3.1, range = 74–86). At the third visit, the mean mini-mental status score was within the normal range (mean = 28.4, SD = 2.0) and the mean Geriatric Depression Scale score indicated no depression (mean = 4.2, SD = 4.1).

Body mass index (BMI; kg/meter^2) significantly increased from visit 1 to visit 2 ($p < 0.001$). The mean BMI for visit 1 was 22.03 (SD = 3.12, range = 17.78–30.35, median = 21.21); mean BMI for visit 2 was 26.01 (SD = 3.62, range = 21.15–34.37, median was 25.28); mean BMI for visit 3 was 26.99 (SD = 3.99, range = 21.00–39.83, median = 26.61).

Changes in sleep apnea

The REML analysis showed no significant differences in either AI or RDI over time, even when controlled for BMI. In other words, there were no significant changes in amount of respiratory disturbance during sleep over the 8.5-year follow-up period. Means for each visit are presented in Table 1.

TABLE 2. *Intervisit Pearson correlations for respiratory disturbance index (p-value)*

	Visit 2	Visit 3
Visit 1	0.75 (0.001)	0.46 (0.05)
Visit 2		0.65 (0.005)

TABLE 3. *Percent with RDI ≥ 15 on visit 1, 2 and 3*

		Visit 2 RDI ≥ 15	Visit 2 RDI < 15	
Visit 1	RDI ≥ 15	26%	17%	n = 10
	RDI < 15	9%	48%	n = 13
		Visit 3 RDI ≥ 15	**Visit 3 RDI < 15**	
Visit 1	RDI ≥ 15	9%	34%	n = 10
	RDI < 15	9%	48%	n = 13
		Visit 3 RDI ≥ 15	**Visit 3 RDI < 15**	
Visit 2	RDI ≥ 15	18%	18%	n = 8
	RDI < 15	5%	59%	n = 14

Pearson correlations were computed between AI for successive visits and RDI for successive visit. Correlations between visit 1 and visit 2 for AI (n = 23) were not significant; however, there was a significant relationship between visit 2 and visit 3 for AI ($r = 0.51$, $p < 0.05$). Correlations between all visits for RDI were significant, indicating that RDI was the more stable measure of long term respiratory disturbance (see Table 2).

For visit 3, night-to-night correlations were computed for those 20 subjects with two nights of recordings. The night-to-night correlation for AI was $r = 0.96$ ($p = 0.001$) but for RDI, $r = 0.65$ ($p = 0.005$).

During visit 1, 38% of the subjects had RDI ≥ 15, but 35% met this criterion at visit 2 and 22% met this criterion at visit 3. Twenty-six percent met the diagnostic level of RDI ≥ 15 on both visit 1 and visit 2 (chi-square = 4.96, $p = 0.026$), and 18% met the criteria on visit 2 and visit 3 (chi-square = 5.32, $p = 0.021$) (Table 3, Fig. 1). The ability to detect sleep apnea at visit 3 based on RDI at visit 1 (i.e. test sensitivity) was only 20%. The specificity (i.e. the ability to detect no sleep apnea) was 85%. The predictive value of RDI ≥ 15 at visit 1 predicting RDI ≥ 15 at visit 3

TABLE 4. *Oximetry values (per night) at visit 3*

	Mean	SD	Median	Range
Number of desaturations (SaO_2)	102.0	121.3	64.5	25.0–439
Mean SaO_2 (%)	93.6	1.9	93.6	90.5–96.5
Lowest SaO_2 (%)	73.3	13.9	78.0	40.0–86.0
Mean low SaO_2 (%)	88.7	2.5	88.9	84.3–92.7
Mean duration (seconds)	46.9	10.9	51.1	28.3–61.6
% time at SaO_2 >90%	92.0	10.7	97.3	71.2–99.9

was 50%. The predictive value of having RDI < 15 at visit 1 and visit 3 was 58%.

Oximetry was only recorded at visit 3, therefore, no longitudinal comparisons can be made for these measures. Table 4, however, shows the oximetry values for descriptive purposes.

Changes in sleep

REML analyses were computed for TST, WASO and number of awakenings counted during the night (Table 5). TST significantly decreased by 1.8 hours from visit 1 to visit 2, but increased again in visit 3 ($p = 0.045$). The number of awakenings decreased by 15–20 awakenings from visit 1 to visit 2 to visit 3 ($p = 0.004$). There were no significant changes in WASO.

Changes in reports of snoring and napping behavior

There was a trend for reports of snoring to increase from snoring 12.7 nights a month (SD = 13, median = 8) to 14.3 times a month (SD = 14.2, median = 8.5) 8.5 years later ($p = 0.062$). Loud snoring increased as a function of time and increases in BMI, from 7.7 times a month (SD = 13.3, median = 0) reported at visit 1 to 11.7 times a month (SD = 13.8, median = 4.25) at visit 2, but then decreased to only 2.7 times a month (SD = 8.3, median = 0) at visit 3 ($p = 0.031$). There were no differences in reports of the number of naps taken either on purpose or inadvertently (i.e. no differences in reports of excessive daytime sleepiness).

DISCUSSION

These longitudinal data indicate that mild SDB did not consistently get significantly better or worse with age, even as weight (as measured by BMI) increased. These data confirm earlier results that indicated that there is great variation in apnea indices over 8.5 years follow-up. Although those with little or no sleep apnea at visit 1 were most likely to have little apnea at subsequent visits, and vice versa, subjects fell into different classes at different follow-up visits. The predictability of a single recording was rather poor after 8.5 years.

The apnea index showed better night-to-night reliability; however, the respiratory disturbance index was more stable over longer time periods. In this sample, the RDI was a more stable predictor of longitudinal severity than apnea index.

Sleep measures also changed over time. Others have noted that total sleep time decreases and number of awakenings increase with age (23,24). Our data indicate that TST did vary with age, but the direction of change varied inconsistently. The number of awakenings, in fact, decreased with age. This lends support to the notion that sleep disturbances seen in the elderly may more likely be consequences of sleep disorders, such as SDB, rather than consequences of aging per se.

We continue to study our sample of 24 elderly. It is imperative to have a complete understanding of the longitudinal course and variability of sleep apnea. With sufficient follow-up, other changes may appear. We will continue to follow and study these elderly individuals to learn more about the longitudinal effects of sleep disordered breathing. In 1992, each will be recontacted for a fourth visit.

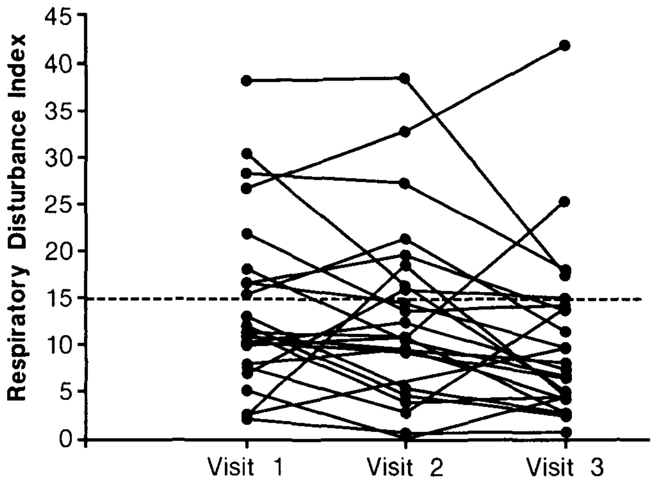

FIG. 1. Respiratory disturbance indices during visits 1, 2 and 3. Individual subjects showed great variability.

Acknowledgements: This study was supported by NIA AG02711, NIA AG08415, NHLBI HL44915 and the Research Service of the Department of Veterans Affairs.

TABLE 5. *Sleep measures by visit*

	Visit 1	Visit 2	Visit 3
TST (p = 0.045)			
Mean	476.6	367.8	448.7
SD	122.9	105.3	79.3
Median	458.3	406.5	467.2
Range	255–673	109–525	205–617
Number of awakenings (p = 0.004)			
Mean	51.6	35.9	15.4
SD	25.2	17.0	9.2
Median	47.0	33.0	13.3
Range	17–105	10–82	2–38
WASO (ns)			
Mean	154.4	86.2	71.6
SD	115.3	74.5	51.2
Median	144.3	76.0	59.9
Range	26–407	26–381	11–23

Abbreviations: TST = total sleep time, SD = standard deviation, WASO = wake after sleep onset, ns = not significant.

REFERENCES

1. Wittig RM, Romaker A, Zorick FJ, Roehrs TZ, Conway WA, Roth T. Night-to-night consistency of apneas during sleep. *Am Rev Respir Dis* 1984;129:244–6.
2. Mosko SS, Dickel MJ, Ashurst J. Night-to-night variability in sleep apnea and sleep in sleep-related periodic leg movements in the elderly. *Sleep* 1988;11:340–8.
3. Aber WR, Block AJ, Hellard DW, Webb WB. Consistency of respiratory measurements from night to night during the sleep of elderly men. *Chest* 1989;96(4):747–51.
4. Kramer M, Silva C. Night to night variability of apnea. *Sleep Res* 1986;15:138.
5. DiPhillipo MA, Goldberg R, Fry JM. Night to night variability of PLMS and apnea in a clinical patient population. *Sleep Res* 1991;20:234.
6. Bliwise DL, Benkert RE, Ingham RH. Sleep apnea nightly variability in aged subjects. *Sleep Res* 1990;19:191.
7. Lee K, Giblin E. Reliability of a one-night diagnostic study for sleep apnea. *Sleep Res* 1982;11:155.
8. Rosenthal LD, Roehrs TA, Roth T. Natural course of sleep apnea: a two-year follow up. In: Kuna ST, Suratt PM, Remmers JE, eds. *Sleep and respiration in aging adults.* New York: Elsevier, 1991:348.
9. Bliwise D, Carskadon M, Carey E, Dement W. Longitudinal development of sleep-related respiratory disturbance in adult humans. *J Gerontol* 1984;39:290–3.
10. Mason WJ, Ancoli-Israel S, Kripke DF. Apnea revisited: a longitudinal follow-up. *Sleep* 1989;12(5):423–9.
11. Phoha RL, Dickel MJ, Mosko SS. Laboratory note. Preliminary longitudinal assessment of sleep in the elderly. *Sleep* 1991;14(4):425–9.
12. Bliwise D, Ingham RH, Nino-Murcia G, Pursley AM, Dement WC. Five-year follow-up of sleep related respiratory disturbance and neuropsychological variables in elderly subjects. *Sleep Res* 1989;18:202.
13. Zorick FJ, Wittig RM, Rosenthal LD, Potts GE, Roehrs TA, Roth T. Natural course of sleep apnea: a 2 year follow-up study. *Sleep Res* 1991;20:354 (Abstract).
14. Ancoli-Israel S, Klauber MR, Fell RL, Parker L, Kenney LA, Willens R. Sleep disordered breathing: preliminary natural history and mortality results. In: Seifert RA, Carlson J, eds. *International perspective on applied psychophysiology.* 1991 (in press).
15. Ancoli-Israel S, Kripke DF, Klauber MR, Mason WJ, Fell R, Kaplan O. Sleep disordered breathing in community-dwelling elderly. *Sleep* 1991;14(6):486–95.
16. Ancoli-Israel S. The use of a modified Respitrace/Medilog portable system in the evaluation of sleep apnea. *J Ambu Monitor* 1988;1(4):267–78.
17. Ancoli-Israel S, Kripke DF, Mason W, Messin S. Comparisons of home sleep recordings and polysomnograms in older adults with sleep disorders. *Sleep* 1981;4(3):283–91.
18. Timms RM, Dawson A, Taft R, Erman MK, Mitler MM. Oxygen saturation by oximetry: analysis by microcomputer. *J Polysomnogr Technol* 1988;Spring:13–21.
19. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1983;17:37–49.
20. Yesavage JA, Brink TL, Rose TL, Adey M. The geriatric depression rating scale. Comparison with other self-report and psychiatric rating scales. In: Crook T, Ferris S, Bartus R, eds. *Assessment in geriatric psychopharmacology.* New Haven, CT: Mark Powles Associates, Inc., 1983:153–67.
21. Folstein MF, Folstein SE, McHugh PR. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98.
22. Laird NM, Ware JH. Random effects models for longitudinal data. *Biometrics* 1982;38:963–74.
23. Kales A, Wilson T, Kales JD, et al. Measurements of all-night sleep in normal elderly persons: effects of aging. *J Am Geriatr Soc* 1967;15:405–14.
24. Reynolds CFIII, Kupfer DJ, Taska LS, Hoch CC, Sewitch DE, Spiker DG. Sleep of healthy seniors: a revisit. *Sleep* 1985;8:20–9.