NEW RESEARCH

JCSM
Journal of Clinical Sleep Medicine

Accuracy and Linearity of Positive Airway Pressure Devices: A Technical Bench Testing Study

Luis Torre-Bouscoulet, M.D.[1,2]; Elodia López-Escárcega, P.T.[1]; José Luis Carrillo-Alduenda, M.D.[1]; Fernando Arredondo-del-Bosque, M.D.[1]; Margarita Reyes-Zúñiga, M.D.[1]; Armando Castorena-Maldonado, M.D.[1]

[1]Sleep Medicine Unit, and [2]Department of Respiratory Physiology, National Institute of Respiratory Diseases, Mexico City, Mexico

SCIENTIFIC INVESTIGATIONS

Study Objectives: To analyze the accuracy and linearity of different CPAP devices outside of the manufacturers' own quality control environment.

Methods: Accuracy (how well readings agree with the gold standard) and linearity were evaluated by comparing programmed pressure to measured CPAP pressure using an instrument established as the gold standard. Comparisons were made centimeter-by-centimeter (linearity) throughout the entire programming spectrum of each device (from 4 to 20 cm H_2O).

Results: A total of 108 CPAP devices were tested (1836 measurements); mean use of the devices was 956 hours. Twenty-two of them were new. The intra-class correlation coefficient (ICC) decreased from 0.97 at pressures programmed between 4 and 10 cm H_2O, to 0.84 at pressures of 16 to 20 cm H_2O. Despite this high ICC, the 95% agreement limit oscillated between −1 and 1 cm H_2O. This same behavior was observed in relation to hours of use: the ICC for readings taken on devices with < 2,000 hours of use was 0.99, while that of the 50 measurements made on devices with > 6,000 hours was 0.97 (the agreement limit oscillated between −1.3 and 2.5 cm H_2O). "Adequate adjustments" were documented in 97% of measurements when the definition was ± 1 cm H_2O of the programmed pressure, but this index of adequate adjustment readings decreased to 85% when the ± 0.5 cm H_2O criterion was applied.

Conclusions: In general, the CPAP devices were accurate and linear throughout the spectrum of programmable pressures; however, strategies to assure short- and long-term equipment reliability are required in conditions of routine use.

Keywords: Sleep apnea, CPAP, accuracy, linearity, quality control, stability

Citation: Torre-Bouscoulet L; López-Escárcega E; Carrillo-Alduenda JL; Arredondo-del-Bosque F; Reyes-Zúñiga M; Castorena-Maldonado A. Accuracy and linearity of positive airway pressure devices: a technical bench testing study. *J Clin Sleep Med* 2010;6(4):369-373.

Obstructive sleep apnea syndrome (OSAS) is a highly prevalent disorder that affects approximately 2% and 4% of women and men, respectively.[1] It is characterized by repeated episodes of complete (apnea) or partial (hypopnea) upper airway obstruction during sleep, events that result in transient oxygen desaturation and are usually terminated by brief arousals.[2] Evidence from well-designed studies indicates that untreated OSAS is independently associated with an increased likelihood of hypertension,[3,4] cardiac events,[5] stroke,[6] motor vehicle accidents,[7] and lower quality of life.[8] Continuous positive airway pressure (CPAP)—treatment of choice for moderate to severe cases—has proven to be an effective strategy that decreases cardiovascular morbidity and mortality.[9] In addition, when used properly, CPAP has been shown to improve quality of life[10] and reduce the risk of traffic and occupational accidents.[11] However, in order to restore upper airway patency and, as a consequence, to achieve the identified benefits associated with the proper use of CPAP,[9] the appropriate pressure level, which is different for each patient, must be precisely guaranteed.[2] A technician usually determines the level of therapeutic CPAP pressure on the basis of a continuous polysomnographic recording,[2] a crucial step because subtherapeutic pressure has not been associated with reduction in the incidence of cardiovascular diseases.[12,13] Therefore, to assure the stability of CPAP devices over time so that they provide the correct level of pressure in long-term use seems particularly important. Thus, the aim of this study was to analyze the accuracy and linearity of different CPAP devices outside the manufacturers´ own quality control environment.

BRIEF SUMMARY

Current Knowledge/Study Rationale: Continuous positive airway pressure is the treatment of choice for patients with obstructive sleep apnea; however, to achieve the indentified benefits associated with this treatment, the appropriate level of pressure must be precisely guaranteed. Thus, to assure the stability of CPAP devices over time so that they provide the correct level of pressure in long-term use seems particularly important.

Study Impact: Although in general, the CPAP devices were accurate and linear throughout the spectrum of programmable pressure, strategies to assure short- and long-term equipment reliability are required. We suggest that, regardless of manufacturers' specifications, accuracy and linearity be verified when the equipment is acquired and at least once a year after that.

MATERIAL AND METHODS

The CPAP devices evaluated include some that were property of Mexico's National Institute of Respiratory Diseases (INER, *Instituto Nacional de Enfermedades Respiratorias*) (both new and in use) and others that belonged to patients who came to the INER's Sleep Clinic regularly. Included were devices providing

Table 1—Concordance between programmed and measured pressures according to level of pressure (upper panel) and hours of use (lower panel)

Programmed pressure (cm H_2O)	n	ICC	95% CI	Difference‡	95% Limit of agreement
4–10	749	0.97*	0.972–0.979	0.090	−0.76 to 0.94
11–15	535	0.90*	0.89–0.92	−0.055	−1.33 to −1.22
16–20	472	0.84*	0.81–0.86	−0.039	−1.53 to 1.45
Total	1756	0.99*	0.991–0.993	0.011	−1.18 to 1.20
Hours of use	**n**	**ICC**	**95% CI**	**Difference‡**	**95% Limit of agreement**
0–2000	1460	0.99*	0.991–0.993	0.007	−1.16 to 1.17
2001–4000	182	0.99*	0.993–0.996	−0.069	−1.08 to 0.94
4001– 6000	49	0.99*	0.996–0.999	0.061	−0.59 to 0.71
> 6000	50	0.97*	0.957– 0.987	0.610	−1.27 to 2.49
Total	1741	0.99*	0.991–0.993	0.018	−1.16 to 1.20

n, number of measurements; ICC, intra-class correlation coefficient; 95% CI, 95% confidence interval; *p < 0.0001, ‡the difference was obtained as follows: programmed pressure minus measured pressure.

fixed-pressure CPAP and others that were self-adjustable and utilized for 7-night titration tests taking place in the home. The self-adjusting models were assessed using their CPAP modality (fixed CPAP). The 2 commercial brands of devices used in testing were: Respironics (Murrysville, PA, USA) and ResMed (Sydney, Australia). Accuracy (how well readings agree with the gold standard) was evaluated by comparing programmed pressure to measured pressure using an instrument established as the gold standard. Comparisons were made centimeter-by-centimeter throughout the entire programming spectrum of each device (from 4 to 20 cm H_2O). Readings were taken once the apparatus had been programmed and after a 1-min interval to allow the pressure to stabilize. Before taking readings, the hose of the CPAP device was connected to a differential pressure gauge (in place of a mask), and the system was checked for leaks. Evaluation of all devices was performed in duplicate by the same person.

This study was approved by the institutional committee on science and bioethics. Because its objective was to analyze the performance of CPAP devices and no clinical or demographic data on users were analyzed, no informed consent was obtained. Patients were not informed as to the condition or "adjustment" of their specific machine, because the biological relevance of the possible differences between programmed and measured pressures is unknown.

The gold standard against which the pressures programmed in the CPAP devices were compared was provided by a Dwyer Puritan Bennet differential pressure gauge (Magnehelic 12-193932-02; Indianapolis, IN) with a 0-to-20 cm H_2O measuring scale and an accuracy of 2% ET. Initial calibration of the differential pressure gauge before commencing the study was conducted at a temperature of 21ºC (70°F) and a relative humidity of 65% using the direct calibration method with a master gauge (TESTO, model 06381447, accuracy = 0.1% ET). The maximum error encountered during initial calibration was −0.22 cm H_2O. At the end of the study, researchers reconfirmed that the gauge was appropriately calibrated to discard the possibility that technical failures in the gold standard might have affected concordance. Calibration of the gauge upon finalizing the study was carried out at 21ºC (70°F) and a relative humidity of 39%. The maximum error of measurement was −0.32 cm H_2O.

To analyze the accuracy of the pressure programmed on the CPAP devices, a concordance analysis between programmed and measured values was conducted, and the intra-class correlation coefficient was calculated (ICC). Concordance is shown by the Bland-Altman modified graph, while linearity was evaluated by calculating the Pearson correlation coefficient, with a p value < 0.05 established as statistically significant. Statistical analysis was conducted using a commercially available program (Stata, release 9.2, StataCorp, College Station, TX, USA).

RESULTS

A total of 108 CPAP devices (23 Respironics Models: REMstar Auto, REMstar Plus with C-Flex, REMstar Pro with C-Flex; and 85 ResMed: Models: S7 Elite, S7 Lightweight, S8 Elite, S8 Escape, S8 AutoSet Spirit) were tested. Seventeen readings were taken with each one (from 4 to 20 cm H_2O), for a total number of 1836 measurements. One device was eliminated from the analysis because the measured pressure exceeded the programmed pressure by > 5 cm H_2O. Upon eliminating that apparatus, the total number of measurements fell to 1819. We also took 63 readings in which the pressure measured exceeded the maximum value of the scale on the gauge used as the gold standard, so those measurements were also discarded. Thus, 1756 measurements were considered useful for concordance analysis. As some of the devices were not equipped with counters indicating hours of use, the concordance analysis that examined this factor was based on 1741 readings. Mean use was 956 hours. Twenty-two of the devices were new.

Table 1 shows the concordance between programmed and measured pressure according to different programmed levels. The ICC decreased from 0.97 at pressures programmed between 4 and 10 cm H_2O, to 0.84 at pressures of 16 to 20 cm H_2O. Despite this high ICC, the 95% agreement limit oscillated between −1 and 1. This same behavior was observed in relation to the hours of use: the ICC for readings taken on devices with < 2,000 hours of use was 0.99, while that of the 50 measure-

ments made on devices with > 6,000 hours was 0.97. Thus, the agreement limit oscillated between −1.3 and 2.5 cm H_2O in the latter group.

Figures 1 and **2** show the correlation and concordance obtained between programmed and measured pressure. Both coefficients, that of intra-class correlation and Pearson correlation, were 0.99; however, the points became more dispersed as the pressure programmed increased (**Figure 2**).

Given that the differences found between programmed pressures and readings may be of little biological importance, we decided to analyze "adjustment." The factor of adjustment was considered "adequate" when the pressure measured was within ± 1 cm H_2O of the programmed pressure, or, more strictly, within ± 0.5 cm H_2O of that pressure. "Low adjustment" occurred when the pressure reading was below that programmed by more than 1, or more than 0.5 cm H_2O, respectively. "High adjustment" occurred when measured pressure was higher than programmed pressure by > 1, or > 0.5 cm H_2O, respectively. **Tables 2** and **3** show the percentage of readings that were judged as adequate, low, and high adjustments, taking into account programmed pressure (**Table 2**) and hours of use (**Table 3**). Adequate adjustments were documented in 97% of measurements when the definition was ± 1 cm H_2O of the programmed pressure; but the index of adequate adjustment readings decreased to 85% when the ± 0.5 cm H_2O criterion was applied. As **Table 1** and **Figure 2** show, the percentage of measurements with adequate adjustment (± 0.5 cm H_2O) decreased as the programmed pressure and hours of use increased, until it reached a level of 78% for programmed pressures between 16 and 20 cm H_2O, and 68% for readings taken with devices with over 6,000 hours of use.

Finally, the concordance between programmed and measured pressures according to the 2 commercial brands tested is shown in **Table 4**.

DISCUSSION

The most important results derived from our study are: (1) in general, regardless of commercial brand, the CPAP devices were accurate and linear throughout the spectrum of programmable pressures; (2) the concordance between programmed and measured pressure, as well as linearity, decrease slightly at pressures above 10 cm H_2O, and after 6,000 hours of use; (3) if we accept as "good" adjustment a variation of ± 1 cm H_2O, then 97% of the readings were adequate in comparison to the programmed pressure; and (4) upon considering a stricter

Figure 1—Correlation between programmed and measured pressures

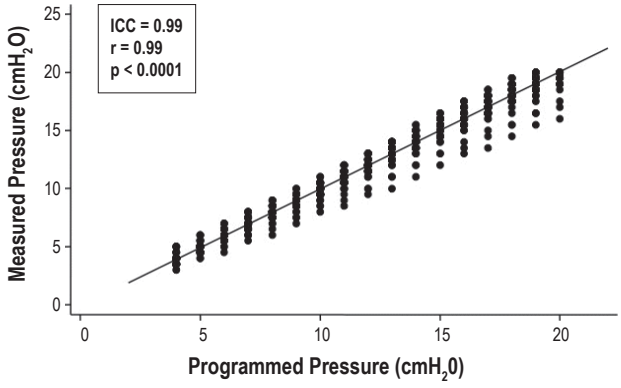
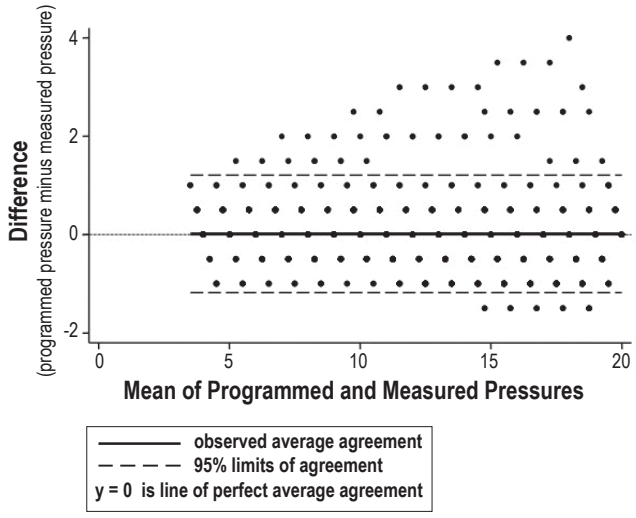


Figure 2—Modified Bland-Altman plot



The difference was calculated as follows: programmed pressure minus measured pressure.

Table 2—Proportion of measurements in which low, adequate, or high adjustments were observed (at ± 0.5 cm H_2O or ± 1 cm H_2O), according to 3 levels of programmed pressure

		Programmed pressure (cm H_2O)			
	Adjustment	4 to 10	11 to 15	16 to 20	Total
± 0.5 cm H_2O	Low	27 (3.6%)	29 (5.4%)	31 (6.6%)	87 (5%)
	Adequate	695 (92.8%)	432 (80.7%)	369 (78.2%)	1496 (85%)
	High	27 (3.6%)	74 (13.9%)	72 (15.2%)	173 (10%)
	Total	749 (100%)	535 (100%)	472 (100%)	1756 (100%)
± 1 cm H_2O	Low	10 (1.3%)	15 (2.8%)	18 (3.8%)	43 (2%)
	Adequate	739 (98.7%)	515 (96.2%)	444 (94%)	1698 (97%)
	High	0 (0%)	5 (1%)	10 (2.2%)	15 (1%)
	Total	749 (100%)	535 (100%)	472 (100%)	1756 (100%)

Table 3—Proportion of measurements in which low, adequate, or high adjustments were observed (at ± 0.5 cm H_2O or ± 1 cm H_2O) according to hours of equipment use

		Hours of use				
	Adjustment	< 2000	2000-4000	4000-6000	> 6000	Total
± 0.5 cm H_2O	Low	64 (4%)	7 (4%)	0 (0%)	16 (32%)	87 (5%)
	Adequate	1250 (86%)	155 (85%)	49 (100%)	34 (68%)	1488 (85%)
	High	146 (10%)	20 (11%)	0 (0%)	0 (0%)	166 (10%)
	Total	1460 (100%)	182 (100%)	49 (100%)	50 (100%)	1741 (100%)
± 1 cm H_2O	Low	30 (2%)	0 (0%)	0 (0%)	13 (26%)	43 (2%)
	Adequate	1419 (97%)	182 (100%)	49 (100%)	37 (74%)	1687 (97%)
	High	11 (1%)	0 (0%)	0 (0%)	0 (0%)	11 (1%)
	Total	1460 (100%)	182 (100%)	49 (100%)	50 (100%)	1741 (100%)

Table 4—Concordance between programmed and measured pressures according to commercial brand

Brand	n	ICC	95% CI	Difference‡	95% Limit of agreement
ResMed	1365	0.99*	0.995–0.996	−0.185	−1.04 to 0.675
Respironics	391	0.98*	0.978–0.985	0.695	−0.537 to 1.928

n, number of measurements; ICC, intra-class correlation coefficient; 95% CI, 95% confidence interval; *p < 0.0001, ‡the difference was obtained as follows: programmed pressure minus measured pressure.

criterion of adjustment (± 0.5 cm H_2O), 85% of readings were judged adequate.

Verification of accuracy, precision, and linearity of the devices should form part of the quality control programs at CPAP clinics, as this would make it possible to assure that during routine use (and beyond the manufacturer's specifications), the programmed pressure will in fact be that which the patient receives. The importance of verifying the stability of other diagnostic and therapeutic devices has been demonstrated in earlier studies.[14]

Though the biological importance of the differences found between programmed and measured pressure is unknown, we believe intuitively that the pressure provided to the patient should not oscillate by more than 1 cm H_2O in relation to the programmed pressure, and that devices with an error above 1 cm H_2O should not be utilized. Conceptually speaking, if the pressure provided to the patient is below that programmed, it could diminish the effectiveness of CPAP in preventing the collapse of the upper airway; while higher pressures, could affect acceptance of and adherence to CPAP, or induce central apneas, especially in susceptible subjects.[15] Other studies have shown that suboptimal CPAP pressures are ineffective in modifying response parameters to treatment, including markers of cardiovascular risk, somnolence, and quality of life.[12]

On the basis of our results, we propose that the accuracy and linearity of CPAP devices be verified upon purchase and then after every 2,000 hours of use, which would correspond to approximately one year. In those cases in which devices have over 6,000 hours of use (approximately 3 years), or the patient requires more than 12 cm H_2O of therapeutic CPAP pressure, we suggest verifying calibration every 6 months.

Our study has limitations. First, the devices evaluated were from just two commercial brands, thus we cannot affirm that the results observed in this study are applicable to other apparatuses. Second, on the basis of this study, we cannot discern if the differences found between programmed and measured pressure, though minimal, might have clinical implications. Moreover, our results do not allow us to assure whether the tidal volume would increase the differences among programmed, measured, and supplied pressures. Finally, the maximum pressure measured by the gauge used as the gold standard was 20 cm H_2O, so pressure readings above this value were not registered and, therefore, not analyzed. Nonetheless, this is to our knowledge the first study that explores the accuracy and linearity of CPAP devices beyond the manufacturers' own quality control programs and in conditions of routine use.

In conclusion, although the observed performance of the CPAP apparatuses was adequate, strategies to assure short- and long-term equipment reliability are required. We suggest that, regardless of manufacturers' specifications, accuracy and linearity be verified when the equipment is acquired and at least once a year after that.

REFERENCES

1. Young T, Palta M, Dempsey J, et al. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;328:1230-5.
2. American Academy of Sleep Medicine. International classification of sleep disorders, 2nd ed: diagnostic and coding manual. Westchester, IL: American Academy of Sleep Medicine, 2005.
3. Nieto FJ, Young TB, Lind BK, et al. Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. Sleep Heart Health Study. *JAMA* 2000;283:1829-36.
4. Peppard PE, Young T, Palta M, et al. Prospective study of the association between sleep-disordered breathing and hypertension. *N Engl J Med* 2000;342:1378-84.
5. Peker Y, Carlson J, Hedner J. Increased incidence of coronary artery disease in sleep apnoea: a long-term follow-up. *Eur Respir J* 2006;28:596-602.

6. Yaggi HK, Concato J, Kernan WN, et al. Obstructive sleep apnea as a risk factor for stroke and death. *N Engl J Med* 2005;353:2034-41.
7. Teran-Santos J, Jimenez-Gomez A, Cordero-Guevara J. The association between sleep apnea and the risk of traffic accidents. Cooperative Group Burgos-Santander. *N Engl J Med* 1999;340:847-51.
8. Baldwin CM, Griffith KA, Nieto FJ, O'Connor GT, Walsleben JA, Redline S. The association of sleep-disordered breathing and sleep symptoms with quality of life in the Sleep Heart Health Study. *Sleep* 2001;24:96-105.
9. Marin JM, Carrizo SJ, Vicente E, Agusti AG. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet* 2005;365:1046-53.
10. Kawahara S, Akashiba T, Akahoshi T, et al. Nasal CPAP improves the quality of life and lessens the depressive symptoms in patients with obstructive sleep apnea syndrome. *Intern Med* 2005;44:422-7.
11. Findley L, Smith C, Hooper J, et al. Treatment with nasal CPAP decreases automobile accidents in patients with sleep apnea. *Am J Respir Crit Care Med* 2000;161:857-9.
12. Becker HF, Jerrentrup A, Ploch T, et al. Effect of nasal continuous positive airway pressure treatment on blood pressure in patients with obstructive sleep apnea. *Circulation* 2003;107:68-73.
13. Peker Y, Hedner J, Norum J, et al. Increased incidence of cardiovascular disease in middle-aged men with obstructive sleep apnea: a 7-year follow-up. *Am J Respir Crit Care Med* 2002;166:159-65.
14. Perez-Padilla R, Vazquez-Garcia JC, Marquez MN, et al. The long-term stability of portable spirometers used in a multinational study of the prevalence of chronic obstructive pulmonary disease. *Respir Care* 2006;51:1167-71.
15. Pusalavidyasagar SS, Olson EJ, Gay PC, et al. Treatment of complex sleep apnea syndrome: a retrospective comparative review. *Sleep Med* 2006;7:474-9.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Roberto Mejía-Alfaro for his critical reading of the manuscript and Mr. Paul Kersey for correcting language use.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication November, 2009
Submitted in final revised form February, 2010
Accepted for publication February, 2010
Address correspondence to: Luis Torre-Bouscoulet, M.D., Departamento de Fisiología Respiratoria y Clínica de Trastornos Respiratorios del Dormir, Instituto Nacional de Enfermedades Respiratorias, Tlalpan 4502, Colonia Sección XVI, Delegación Tlalpan CP 14080, México, D.F.; Tel and Fax: (52) (55) 56654748; E-mail: luistorreb2001@yahoo.com.mx

DISCLOSURE STATEMENT

This was not an industry support study. The authors have indicated no financial conflicts of interest.