ANOVA OF FUNCTIONAL DATA AMONG SLEEP APNEA PATIENTS

NZM Saat[1], AA Jemain[2] & SHA Al-Mashoor[3]

[1]Programme of Biomedical Science, School of Science Diagnostic and Health Science, Faculty of Health, Universiti Kebangsaan Malaysia

[2]School of Mathematical Science, Faculty of Science and Technology, Universiti Kebangsaan Malaysia

[3]Faculty of Medicine, University Technology MARA, Level 11 Selayang Hospital, Selangor, Malaysia

ABSTRACT

Apnea is one of the sleep disorder and many study relates apnea with gender,age and BMI of the patients. Few study indicates that apnea occurred in REM and NREM sleep stages.This study involves a sleep record from apnea patients in Hospital Sarawak, Malaysia. The objective of the study is to find the homogenous of REM sleep between patients using the Gini index. Furthermore the differences of mean functional of REM according to gender, age and bmi. The results shows that the Gini index is between 0.16 until 0.64. Meanwhile from the analysis variance of functional data indicates that rate of REM was not differed according to age group, BMI and RDI.In conclusion,the distribution of REM was homogeneous among patients.

Keywords: *REM;apnea; Gini index;functional data;BMI.*

1. INTRODUCTION

Apnea is a paused in breathing during sleep (Tsai et al.1999) . Patients that is di- agnosed with apnea problem increased since 1980's(Guilleminault et al.1973).The prevalence of apnea differ according to race and gender. Study by Redline et al.2004 showed that the prevalence of apnea is higher among blacks compared to whites(Meetze et al., 2002). In Korea study by Kim et al. 2004 showed that prevalence of apnea is higher among males which is 27% compared to females which is 16%.In most of study of apnea the apnea is related to hypertension and obesity problem. Study by Meetze et al. 2002 showed that the prevalence of apnea is higher among black and obese patients.

Sleep apnea prevalence is increasing rapidly for example in Young et al., 1993. The apnea hypopnea index(AHI) is being used to indicate the severity of apnea. The category of severity of apnea using the AHI is <5, <15 and < 30 (Duran et al. 2001). AHI is defined as number of apnea and hypopnea per hour (Tsai et al.1999). Even though the higher AHI indicating the more seriousness of apnea for the particular patient, further study is needed , in order to determine the proportional distribution of apnea for a period of time, say 10% period of sleep.This information may shed some light on understanding the behaviour occurrence of apnea among adults.

In order to study the disproportional or the inequality occurrence of apnea Lorenz curve, Gini index and Run test are used. Lorenz curve and Gini index are widely being used in economy sector. For example study of inequality of poverty distribution in Malaysia(Anand , 1977).Study by Damgaard and Weiner(2000) and He et al. (2005) used Lorenz curve to determine the inequality of plant size. In medical, the Gini index have been used to describe the inequality of healthcare among low income patients(Anderson et al. 2002).

Study on functional data has arise since 1990's. FDA mostly using continous data especially in study involves medical(Manteiga & Vieu 2007). The suitable data for FDA is a time series data or longitudinal data that is not equal space(Ramsay & Dalzall 1991).Many study applied the use of FDA and transformed to smooth function has been discussed by Ramsay & Silverman 2006. Most studies using statistical analysis such as analysis of variance (Cuevas et al. 2006); principal component(Aguilera et al. 2006)to analyze functional data.

The current study wants to determine the variation of occurrence of REM according to gender, age group and respiratory disturbance index(RDI) in application to analysis of variance for functional data(Cuevas et al., 2006).REM is a sleep stages. Many studies , for example by Siddiqui et al. 2006 that relates REM and age, gender and BMI. Furthermore study by Redline et al. 2004 also describes the differences of apnea during REM and Non REM sleep stages. The total REM for each patients is not the same, because the sleeping period for each patients varies. Although it is beneficial to study the relationship between REM and covariates, this might overlook the effect of REM for difference period hour of sleep, as every patients has difference total hour of sleep.In general the func- tion at which the point of time was observed for example every 10% period of sleep, or every 25% period of sleep may varies according to patients.The repeated measurement of REM can produce different curves and its vary according to patients. Some analysis variance(ANOVA) may able to find out the differences of mean duration of REM between age group, but the data of REM was measured for every 30 seconds during sleep , but ANOVA didn't take into account the

time space between occurrence of REM.Therefore it is more reliable to adapt FDA as the procedure consider the time space between REM.

2. METHODS

The study was conducted between 1996 and 1998 in Hospital Sarawak, Malaysia . Suprisingly there was not much study on severity, prevalence of apnea in Malaysia. Study in Malaysia, by Kamil and co-workers in 2007, indicated that 7% of 1611 patients was suspected to have apnea.In statistical model study by Saat et al.(2010) using data from MIT-BIH indicating that the the appropriate order of Markov chain was three ,regardless of sleep stages. In this study a apnea events was consist of obstructive sleep apnea(OSA), mixed apnea(MA) and hypopnea. It is because most of the apnea events among patients was hypopnea. The mean duration of apnea during REM and NREM sleep stages, was consist of duration of apnea during OSA,MA and hypopnea.

2.1 Lorenz curve and Gini index

By using Lorenz curve, the rate of apnea for every 10% period of time was sort from smallest to the highest. Then the cumulative of rate of apnea was plotted (y axis) against cumulative proportion of time(x axis). The time indicating the period of sleep and was in percentage of 10% until 100%. Consider a n ordered rate of apnea for duration of sleep until time t , such that zi is the rate of apnea, and z_1 , z_2, .., z_t, i = 1, ..., t, the Lorenz curve is the joined of the rate z_i and become a points of (r/t, L_r/L_t), r was the sum of apnea for the period 10%, 20%, , ...100% and t was the sum duration of sleep in epoch, where L_0 = 0 and Lt was the sum of apnea during time i. The Lorenz curve can be written as follows

$$F(z) = Pb(Z \leq z)$$

$$L = \sum_{z_i \leq z}^{t} F(Z = z) = \sum_{z_i \leq z}^{t} f(z_i)$$

All cumulative rate of apnea that is differ than the horizontal line of 45^0 indicat ing that the distribution of the rate of apnea was inequal. Mean while, by using Lorenz curve it is difficult to indicate if the occurrence of apnea was higher in the beginning, middle or end of sleep. Therefore the cumulative rate of apnea was plot against the time . The cumulative rate of apnea will give information if the line of the rate was high in the beginning of sleep, then the line will be above the horizontal 45^0. Furthermore, if the line is below the horizontal 45^0, indicating that the rate of apnea happen mostly at the end of sleeping period.

Gini coefficient was also used to measure the inequality or how much variation of REM between 10% period of sleep for each patients. It is calculated as the sum of difference between the total area under the uniform distribution which is 45^0 line and the Lorenz curve. The equation of Gini coefficient is given in equation (1) where x is the ordered sum of apnea event for every 10% period of sleep, i is the rank based on the length of data, μ is the mean of apnea event for particular patient and n is the length of apnea event and for this study it is 10 because the data was group into 10 period of sleep duration. The minimum value of Gini coefficient is zero indicating equality of apnea distribution for every 10 % period of sleep and the maximum value is one indicating the distributions is unequal for every 10 % period of sleep(Andersen et al. 2002).

$$G = \sum_{i=1}^{n} \frac{2i - n - 1}{n^2 \mu} \qquad \textbf{(1)}$$

2.2 Analysis Variance of the Functional Data Analysis

In the present paper the Analysis Variance(ANOVA) of the FDA has been used to find the variation or similarity between the REM sleep stages among gender, age group and bmi. Following Cuevas et al.(2006) comparing the k group of function R_{ij} with j = 1, .., n for group i = 1, .., k, the F statistic can be written as

$$F = \frac{\sum_{i=1}^{k} n_i \left\| \hat{X}_{i.} - \hat{X}_{..} \right\| / (k-1)}{\sum_{i=1}^{k} \left\| \hat{X}_{ij} - \hat{X}_{i.}{}^2 \right\| / (n-k)} \qquad (2)$$

The null hypothesis was there is no mean difference of REM between age group and bmi. If the null hypothesis was rejected , indicating that there was mean difference of REM sleep stages. In the present paper permutation technique(Good 2000) will be used to identify the differences between group. The permutation steps is given below

Step 1

Calculate the mean Rij and apply to the value of F statistic using formula(2)

Step 2

Resample the data ,for example for time point 10%(10,20,30,..100), after resample the time point is (20, 40,10,70,..).

Step 3

Calculate the new value of F statistic using formula (2)

Step 4

Repeat step 2 and 3 for Z times, Z=10000 times. Then compare the original value of F stat with the new value of F stat.

Step 5

The probability of rejecting Ho is $p = (F \text{ original} \geq F \text{ new})/Z$

3. RESULTS

The results was presented into two section, first part was the distribution of the REM. In second part was the results from the ANOVA of FDA.The range of Gini index can be from zero until one. When the value of Gini near zero indicating that the distribution of REM is equal foir every 10% period of sleep. In this study we choose 10% period of sleep for simplicity purposes. Figure 1 showed that the distribution of REM sleep stage for every 10% period of sleep was inequal between patients. However, the results of Gini index indicating that the inequality of the REM distribution is moderate, as all the subjects has Gini index below 0.7.However patients 22 has the highest inequality of distribution of REM which is 0.636. Meanwhile patients 5 and 6 has low inequality of distribution of REM which is 0.173 and 0.174. This indicates that the distribution of REM was almost similar for every 10% period of sleep.

In order to find difference between age, bmi and RDI group ANOVA with application of FDA was used. Besides that, the present study also finding the mean differences of interaction factor. The significance mean difference was indicated by calculating the probability of F the original F statistic value is more than the F statistic value from the permutation test.

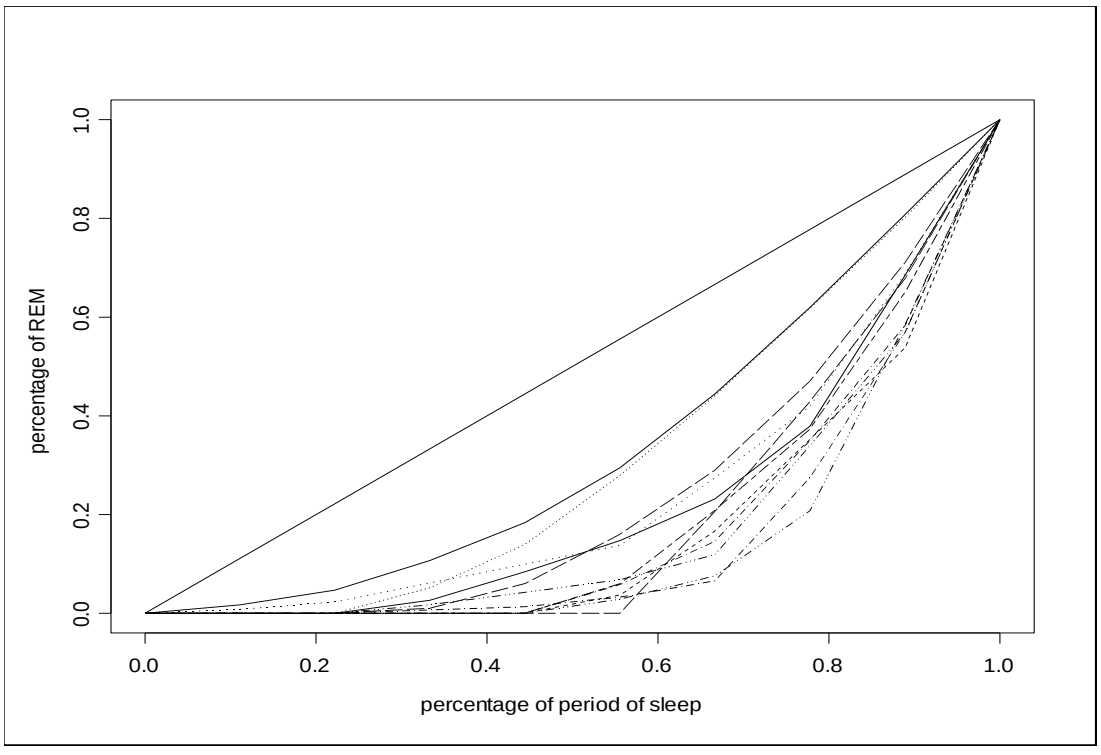

Figure 1:Lorenze curve of the percentage of REM for every 10% period of sleep

RDI	Gini index	Percentile			
		2.5%	5%	95%	97.5%
< 5	0.285	0.177	0.181	0.334	0.373
	0.424	0.099	0.106	0.219	0.235
	0.46	0.089	0.099	0.202	0.212
	0.174	0.271	0.073	0.156	0.163
	0.496	0.275	0.279	0.332	0.3351
	0.481	0.268	0.272	0.325	0.33
	0.585	0.269	0.276	0.331	0.336
	0.489	0.275	0.28	0.329	0.331
	0.471	0.276	0.279	0.329	0.335
	0.568	0.255	0.267	0.354	0.355
	0.4	0.268	0.275	0.3331	0.338
	0.434	0.258	0.262	0.352	0.356
	0.224	0.246	0.2568	0.333	0.337
	0.636	0.259	0.265	0.344	0.348
5 − 15	0.173	0.267	0.271	0.332	0.334
16 − 30	0.355	0.262	0.268	0.341	0.348
	0.24	0.271	0.278	0.333	0.339
	0.606	0.241	0.246	0.349	0.359
	0.564	0.262	0.267	0.336	0.342
	0.454	0.262	0.271	0.346	0.35
	0.491	0.258	0.2691	0.336	0.344
	0.166	0.248	0.255	0.362	0.372

Table 1: Permutation of REM for 10000 simulation of Gini index

		Percentile				
Variable	F_n	2.5%	5%	95%	97.5%	Sig
Age	0.99	0.37	0.42	1.17	1.268	>0.05
BMI	0.883	0.39	0.43	1.02	1.09	>0.05
RDI	0.85	0.35	0.39	1.21	1.31	>0.05
Gender x Age	0.75	0.4	0.43	0.78	0.81	>0.05
Gender x BMI	0.54	0.28	0.30	0.54	0.57	>0.05
Gender x RDI	0.72	0.33	0.36	0.72	0.76	>0.05
Age x BMI	0.34	0.22	0.23	0.39	0.41	>0.05
Agex RDI	0.44	0.29	0.31	0.54	0.57	>0.05
BMI x RDI	0.49	0.32	0.34	0.59	0.61	>0.05

Table 2: The percentile of Fn to find mean differences between factors

The result indicates that there was no significance mean difference between age groups, gender , BMI and RDI. This indicated that the pattern of REM sleep stages is similar for younger and older age group. In addition the pattern of REM sleep stages is similar between male and female subject and patients with lower and higher BMI.

4. DISCUSSION AND CONCLUSION

This paper discuss the distribution of REM during sleep. The Lorenz curve,Gini index and cumulative graph has been used to describe the inequality distribution of REM. In this study the Gini index between patients varies , but it is a moderate inequality. Therefore the distribution of REM was almost similar during the beginning, middle and end of sleep. However study by Goh et al.(2000) describe that REM sleep stages was mostly at the end of sleep. This study is persistent with study by Belyavin (1992) indicated that the REM varies during three period of REM. It is because the interuptions during sleep which is arousal that influence the change of transition from REM to nonREM sleep stages.

Study on distribution of REM has been done by many researcher before. Findley et al.(1985) discribed that apnea was mostly among during Non-REM sleep stages , among older age group and obesity patients. However study by Siddiqui et al.(2006) found that apnea was worst during NREM and obesity patients. Study by Bliwise et al. (1988) indicated that duration of REM was lower among older age group. Data of REM was non-periodic and was repeated measures. It is suitable to be analyzed as functional data. In this study we want to determine the difference of mean occurrence of REM among age group, bmi and RDI. In contrast from the study the results indicated that the mean REM was not significantly difference between age, bmi and RDI. The similar result also shown for interaction term.

Study by Goncalves et al.(2004) showed that the sleep disruptions such as narcolepsy influence the REM. In addition, the percentage of REM was higher among patients that have higher AHI. In contrast, from the results of

anova functional data in this study, there was no mean difference of REM distribution between RDI (RDI<5, 5-15 and RDI>15). Study by Bixler et al. (1998) indicated that REM was not associated with excessive daytime sleepiness(EDS). Patients that has EDS, have a low quality of sleep. This means that eventhough the patients have the EDS the distribution of REM can be high or low.In this study we found that the occurrence of REM was independent with gender. This is similar to the study by Rubio et al.(2005) indicated that there was no significance difference of REM between male and female patients, eventhough REM was higher in male among patients with RDI<15.

In conclusion,the current study showed that the distribution of REM sleep stages was almost similar for every 10% period of sleep. However the distribution of REM were difference depending on age and gender. In the future in order to understand more the distribution off REM sleep stages,it is suggested that the occurrence of REM was studied with other factors such as hypertension, diabetes, smoking habits and EDS.

5. ACKNOWLEDGEMENT

The authors wish to thank Ministry of Higher Education and Universiti Kebangsaan Malaysia for the research grant ERGS/1/2011/SKK/UKM/03/12.

REFERENCES

[1]. Aguilera, A.M.., Escabias M & Valderrama MJ. 2006. Using principal components for estimating logistic regression with high-dimensional multicollinear data.*Computational Statistics & Data Analysis* 50 : 1905-1924
[2]. Anand, S., 1977. *Aspects of poverty in Malaysia review of income and wealth*.23: 1-16
[3]. Andersen RM, Yu H, Wyn R, Davidson PL, Brown ER, Teleki S.2002.Access to medical care for low-income persons: how do communities make a differ enceed? *Care Res Rev* 59(4):384-411.
[4]. Belyavin, A.J. 1992.The duration of REM sleep episodes in normal sleep. J. Sleep Rex . 1: 128-131
[5]. Bliwise,D. L., Bliwise,N. G. , Partinen,M., Pursley,A. M. & Dement,W.C.1988. Sleep apnea and mortality in an aged cohort. *Am J Public Health* 78(5): 544-547.
[6]. Bixler,E.O., Vgontzas,A.N., Have,T. T., Tyson,K. & Kales,A. 1998. Effects of age on sleep apnea in men: I. Prevalence and severity .*Am. J. Respir. Crit. Care Med.* 157(1) : 144-148.
[7]. Cuevas,A., Febrero,M. & Fraiman,R.2006 . *Computational Statistics & Data Analysis* 51(2): 1063-1074
[8]. Damgaard, C. & Weiner, J. 2000.Describing inequality in plant size or fecundity. *Ecology*, 81(4): 1139-1142
[9]. Duran J., Esnaola S., Rubio, R., & Iztueta, A. 2001. Obstructive Sleep ApneaHypopnea and Related Clinical Features in a Population-based Sample of Subjects Aged 30 to 70 Yr. *Am J Respir Crit Care Med* . 163: 685-689
[10]. Findley LJ, Wilhoit SC, Suratt PM. 1985Apnea duration and hypoxemia during REM sleep in patients with obstructive sleep apnea. *Chest* ; 87(4):432-436.
[11]. Good P.I.2000.Permutation tests: A practical guide to resampling methods for testing hypotheses. 2nd edition.Springer. New York.
[12]. Goncalves, M.A., Paiva, T., Ramos E., & Guilleminault C. 2004. Obstructive Sleep Apnea Syndrome,Sleepiness, and Quality of Life. *Chest*. 125:2091-2096
[13]. Goldberger, A.L.,Amaral, L.A.N., Glass, L.,Hausdorff, J.M., Ivanov, P.C.H., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K.& Stanley, H.E. 2000.PhysioBank, PhysioToolkit and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215-e220 [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215
[14]. Goh, D.Y.T., Galster, P., & Marcus, C.P. 2000. Sleep Architecture and Respiratory Disturbances in Children with Obstructive Sleep Apnea *Am J Respir Crit Care Med* 162: 682-686
[15]. He, Z., Z. Ma, K.M. Brown and J.P. Lynch, 2005. Assessment of inequality of root hair density in arabidopsis thaliana using the gini coefficient: A close look at the effect of phosphorus and its interaction with ethylene. *Ann. Botan* y, 95: 287-293.
[16]. Kamil, M.A., Teng,C.L.& Hassan, S.S. 2007 . Snoring and breathing pauses during sleep in the Malaysian population. *Respirology* 12(3): 375 – 380
[17]. Kim, J., K. In, J. Kim, S. You and K. Kang et al., 2004. Prevalence of sleep disorder breathing in middle-aged Korean men and women. *Am J. Respiratory Critical Care Med*. 170: 1108-1113.
[18]. Manteigaa, W.G. & Vieub, P.2007.Statistics for Functional Data. *Computational Statistics & Data Analysis* 51 : 4788 – 4792
[19]. Meetze K, Gillespie MB, Lee FS.2002 . Obstructive Sleep Apnea: A Comparison of Black and White Subjects.The Laryngoscope 112(7): 1271-1274
[20]. Ramsay, J.O. & Dalzell C.J. 1991. Some Tools for Functional Data *Analysis. Journal of the Royal Statistical Society. Series B (Methodological)*, 53(3) : 539-572
[21]. Ramsay,J.O. & Silverman, B.W.2002. Applied functional data analysis. Methods and case studies. Springer Series in Statistics
[22]. Redline, S., Kirchner, H. L., Quan, S. F.,Gottlieb, D. J. ,Kapur& V., New- man, A. 2004 . The Effects of Age, Sex, Ethnicity, and Sleep-Disordered Breathing on Sleep Architecture, *Arch. Intern. Med*.164: 406-418.
[23]. Rubio, J.H.,Janssens, J.P., Thierry Rochat,&Emilia Sforza. 2005. Rapid Eye Movement-Related Disordered Breathing.Clinical and Polysomnographic Features *Chest* 128:3350-3357
[24]. Siddiqui, F. , Walters, A., Goldstein,D., Lahey,M. & Desai, H. 2006. Half of patients with obstructive sleep apnea have a higher NREM AHI than REM AHI. *Sleep Medicine* 7(3): 281-285
[25]. Saat, N.Z.M., A.A. Jemain and K. Ibrahim, 2010a. Statistical model of the occurrence of sleep apnea. *J. Applied Sci*., 10: 752-758.
[26]. Tsai A W.H., Flemons W. W., Whitelaw W. A. &. Remmers J. E. 1999.Com- parison of ApneaHypopnea Indices Derived from Different Definitions of Hypopnea . *Am J Respir Crit Care Med* . 159: 43-48
[27]. Young,T., Palta, M., Dempsey,J., Skatrud,J., Weber,S.& Bad, S. 1993. The Occurrence of Sleep-Disordered Breathing among Middle-Aged Adults.*The New England Journal of Medicine* ,328(17):1230-1235