

# Statistical Modelling technique in Forecasting of Palliative Oncotherapy load in Hospitals

Sathian B<sup>1</sup>, Sreedharan J<sup>2</sup>, Sharan K<sup>3</sup>, Baboo N S<sup>4</sup>, Chawla R<sup>5</sup>, Chandrasekharan N<sup>6</sup>, Rajesh E<sup>7</sup>, Baniya R<sup>5</sup>, Shah R K<sup>5</sup>, Dixit S B<sup>8</sup>

<sup>1</sup> Assistant Professor, Department of Community Medicine, Manipal College of Medical Sciences, Pokhara, Nepal

<sup>2</sup> Assistant Director, Research Division, Gulf Medical University, Ajman, United Arab Emirates

<sup>3</sup> Assistant Professor & Unit chief, Radiotherapy Department, Manipal College of Medical Sciences, Pokhara, Nepal

<sup>4</sup> Professor, Department of Physiology, Manipal College of Medical Sciences, Pokhara, Nepal

<sup>5</sup> MBBS Final Year Students, Manipal College of Medical Sciences, Pokhara, Nepal

<sup>6</sup> Assistant Professor, Department of Orthopaedics, Manipal College of Medical Sciences, Pokhara, Nepal

<sup>7</sup> Assistant Professor, School of Behavioural Sciences, MG University, Kottayam, Kerala, India

<sup>8</sup> Professor & Head, Department of Community Medicine, Manipal College of Medical Sciences, Pokhara, Nepal

---

## Original Article

---

### Corresponding Author:

Dr. Brijesh Sathian, Assistant Professor, Department of Community Medicine, Manipal College of Medical Sciences, Pokhara, Nepal.

Email: [drsathian@gmail.com](mailto:drsathian@gmail.com)

---

## Abstract

---

### Background

An increase in cancer incidence in developing countries is expected due to the possible decline of mortality from infectious diseases, improved healthcare, population growth and increasing life expectancy. It can therefore be anticipated that the number of cancer patients requiring palliative treatment will also rise in Nepal, creating a discrepancy between the requirement and delivery of

health care. Estimation of trends is very essential for the planning of future requirements in any such programme.

### Method

This retrospective study analysed the records from the Radiotherapy Department at Manipal Teaching Hospital, Pokhara, Nepal to calculate the number of cancer patients who received palliative treatment with radiotherapy between September 2000 and December 2008. Descriptive statistics was used for the analysis and subsequently, statistical modelling techniques were applied for forecasting of future trends.

### Results

Of 1001 cancer cases, 363 patients received radiotherapy with palliative intent during the study period. Excluding the constant term from the equation, the Logarithmic model was the best fitted, with  $R^2 = 0.727$ ,  $p = 0.002$  for the forecasting of cancer patients receiving palliative treatment. Using this model, the number of cancer cases receiving palliative radiotherapy at the hospital in the year 2015 was

estimated to be 68. Thus, the number of patients requiring radiotherapy with palliative intent shows a pattern of increasing trend over the coming years.

### Conclusion

The data analysed in the present study indicates an increasing future trend of patients requiring palliative therapy at the centre. In the future, a discrepancy may arise between the requirement and capacity of care available to these patients. Government and healthcare agencies of Nepal must ready themselves to promote better strategies for adequate provision to the cancer patients receiving palliative treatment in the coming years.

### Key Words

Statistical Modelling, Palliative Oncotherapy, Forecasting

---

### Background

The World Health Organization defines palliative care as *'an approach that improves the quality of life of patients and their families facing the problems associated with life threatening illness, through the prevention and relief of suffering by the means of early identification, impeccable assessment and treatment of pain and other problems-physical, psychosocial and spiritual.'* Palliative care is the active, holistic care of patients with terminal illness. Management of pain and other symptoms and provision of psychological, social and spiritual support is of paramount importance. The goal of palliative care is achievement of the best quality of life for patients and their families. Many aspects of palliative care are also applicable earlier in the course of the illness in conjunction with other treatments<sup>1</sup>. Based on the intent, treatment of cancers is classified as either curative or palliative. Though palliative care is not just limited to patients with advanced cancers, they form the majority of patients requiring palliation. Palliative anti-cancer treatment is indicated when curative treatment is not possible and cancer has extensively metastasized but where treatments may have sufficient anticancer activity to improve cancer-related symptoms<sup>2</sup>. This form of treatment is given in addition to symptomatic treatment. Those in terminal phase of illness are usually provided only Best Supportive Care (BSC). Jassem et al defined BSC as "treatment administered with the intent to maximize quality of life without a specific anti-neoplastic regimen"<sup>3</sup>. In 2008, 53% of the 12.7 million new cases of cancer and 63% of the 7.6 million cancer deaths worldwide occurred in developing countries. Projections from GLOBOCAN 2008 also suggest that this trend is likely to remain over the next couple of decades<sup>4</sup>. Many of these patients in developed countries would have received palliative treatment, as well as hospice care. The Centre for Disease Control reported that 55,500 cancer patients received palliative care, and accounted for 52.6% of patients receiving palliative care in the United States of America in the year 2000. However, the situation

is likely to be different in developing and under-developed countries, where the availability of services is far less<sup>5</sup>. A population based cancer registry [the database of cancer cases occurring in a defined population] assists not only in knowing the burden of the disease but also in planning and evaluating health services for the prevention, diagnosis and treatment of the disease<sup>6</sup>. Patients with advanced stage of cancer respond poorly to the treatment and therefore, an early diagnosis of cancer helps for possible cure of the disease. Nepal lacks a population based cancer registry. However, a hospital based-cancer registry's initial report (data taken from 7 hospitals) for the year 2005 mentions a total of 4,397 cancer cases with 2,340 females and 2,057 males<sup>7</sup>. The precise forecasting of cancer trends and the estimation of future incidence rates help in pre-planning and allows for better resource allocation, which is invaluable for public health services within the present situation of Nepal. One approach to study the incidence data on cancer is by analyzing of trends and variations and creation of the best statistical prediction models<sup>8</sup>. The objective of this study was to estimate the numbers, and trends of cancer patients receiving palliative treatment in the Department of Radiotherapy at Manipal Teaching Hospital, Pokhara, Nepal.

### Method

Forecasting of cancer cases for palliative treatment was attempted using the data retrieved from the radiotherapy treatment records of patients treated at Manipal Teaching Hospital between September 2000 and December 2008. The variables collected were age, gender, cancer cases, type of treatment being received (curative vs. palliative), information regarding unplanned interruptions and compliance to the treatment. Out of 1011 cancer cases, 10 lacked relevant data and hence only 1001 cases were analysed in this study. The cancer cases were categorized into groups, based on the site of lesion: brain, head and neck, breast, lung, cervix, gastrointestinal tract (GIT), hepatobiliary, urogenital, and gynaecological cancers. Head and neck cancers included cancers of oral cavity, tongue, oropharynx, larynx, buccal mucosa, maxilla, paranasal sinuses, parotid glands and nasopharyngeal carcinomas. Gynaecological cancers included ovarian, uterine and vulvar/vaginal cancers. Cervix cancer was categorized as a separate category due to its high incidence. Cancers not falling into above mentioned categories were put into 'others'. Approval for the study was obtained from the Institutional Research Ethical Committee. Excel 2003, R 2.8.0, Statistical Package for the Social Sciences (SPSS) for Windows Version 16.0 (SPSS Inc; Chicago, IL, USA) and EPI Info 3.5.1 Windows version were used for statistical modelling and prediction. Linear, Logarithmic, Inverse, Quadratic, Cubic, Compound, Power, Exponential, and Growth curve fitting models were tested. F-test was used for selecting the best fitting curve for hypothesis testing. P-value  $\leq 0.05$  (two sided) was taken as statistically significant. Similarly,  $R^2$  value  $> 0.80$  was taken

significant for more precise prediction, where  $R^2$  is the correlation of the contribution of years (independent variables) in predicting cancer cases (dependent variables)<sup>9</sup>. The decision regarding the selection of a suitable prediction approach is governed by the relative performance of the models for monitoring and prediction. It should also adequately interpret the phenomenon under study. The Logarithmic model selected here was able to closely fit curves for estimated and observed palliative intent cancer cases (Fig 1, Fig 2). While building a model, the extremities (maximums and minimums) play a crucial role. If the points are scattered more, the curve tries to adjust with maximum number of observed points. The Logarithmic model is represented by the equation  $Y = m_0 + m_1 \cdot \log(x)$ , where  $m_0$  is the constant term and  $m_1$  is coefficient terms<sup>10,11</sup>. Without the constant term, the equation of this model is  $Y = m_1 \cdot \log(x)$ .

**Model validation**

Residuals are used to validate the models. Minimum and maximum residual values play important role in statistical forecasting. The criterion to find out the best validated model is one having least residual.

**Results**

Of 1001 cancer cases, 363 patients received radiotherapy with palliative intent for their treatment during the study period. 37% of all the palliative intent cancer cases were of lung. Out of the total 209 lung cancer patients, 63.6% received radiotherapy with palliative intent.

**Table 1: Distribution of patients receiving palliative radiotherapy according to the site of primary tumour**

Case	Total Cancer cases	Palliative cancer cases	%	Confidence Interval
Brain	17	2	11.8	(0, 27)
Breast	70	14	20	(10.6,29.4)
Cervix	159	19	11.9	(6.9, 17)
GIT	88	48	54.5	(44, 65)
Other	17	5	29.4	(7.8, 51)
Gynaecological malignancies				
Head & Neck	242	25	10.3	(6.5, 14)
Hepatobiliary	4	3	75	(32.6, 100)
Lung	209	133	63.6	(57, 70)
Urogenital	57	32	56.1	(43.3, 69)
Others	138	82	59.4	(51.2,67.6)
<b>Total</b>	<b>1001</b>	<b>363</b>	<b>36.2</b>	<b>(33.3,39.2)</b>

The number of patients receiving radiotherapy with palliative intent showed a pattern of increasing trend. The data were modelled using Curve Fitting method.

Table 2 depicts the model summary and parameter estimates including the constant term for different models,

when the constant term was included, the p values were >0.05 in all the models, and none of the models were best fitted.

**Table 2 Model Summary and Parameter Estimates including the constant term for different models**

Model	Model summary	
	R <sup>2</sup> value	P value
Linear	0.001	0.933
Logarithmic	0.069	0.494
Inverse	0.198	0.230
Quadratic	0.597	0.066
Cubic	0.664	0.116
Compound	0.017	0.737
Power	0.143	0.315
S	0.347	0.095
Growth	0.017	0.737
Exponential	0.017	0.737

**Table 3 Model Summary and Parameter Estimates excluding the constant term for different models**

Model	Model summary		Parameter Estimates		
	R <sup>2</sup> value	P value	m <sub>1</sub>	m <sub>2</sub>	m <sub>3</sub>
Linear	0.632	0.006	6.425		
Logarithmic	0.727	0.002	24.614		
Inverse	0.292	0.107	59.404		
Quadratic	0.909	0.001	23.034	-2.34	
Cubic	0.909	0.002	24.216	-2.76	0.033
Compound	0.782	0.001	1.756		
Power	0.841	0.000	2.085		
S	0.459	0.031	5.866		
Growth	0.782	0.001	0.563		
Exponential	0.782	0.001	0.563		

After excluding the constant term, all curves (except for inverse curve) fitted well with the data. Table 3 displays model summary and parameter estimates excluding the constant term for different models and Fig. 1 shows fitted curves for observed palliative intent cancer cases. Excluding the constant term from the equation, the Logarithmic model was the best fitted with  $R^2 = 0.727$ ,  $p = 0.002$  for the forecasting of cancer patients receiving palliative treatment.

The equation for the logarithmic model for Palliative cancer cases is

$$Y = 24.61 \cdot \log(x) \text{ --- (1)}$$

Using the equation (1), Palliative cancer cases were estimated. Table 4 shows the observed and estimated number of Palliative cancer cases up to the year 2008, and estimated number of cancer cases up to the year 2015.

Figure 1 Fitted curves for observed palliative intent cancer cases

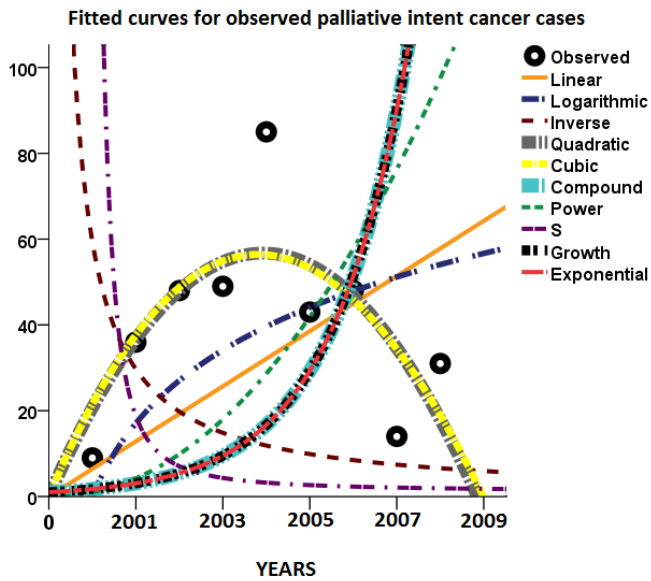


Table 4 Observed and estimated Palliative cancer cases

Year	Observed	Estimated
2000	9	-
2001	36	17
2002	48	27
2003	49	34
2004	85	40
2005	43	44
2006	48	48
2007	14	51
2008	31	54
2009		57
2010		59
2011		61
2012		63
2013		65
2014		67
2015		68

**Model validation**

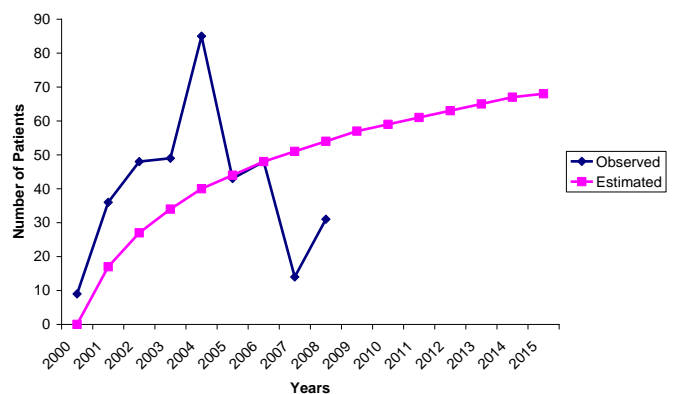
Table 5 depicts the validation of all models. Negative values show under forecasting and positive values show over forecasting. Quadratic and cubic models have the least residual values. Based on the mathematical equation, it is the curve's characteristic to fit in maximum observed values. However, because of the scattered values and influence of minimum and maximum, in this model it has lead to the loss of its predictive nature as seen for 2009 where frequency reached zero (figure 1). Compound, power, growth and exponential models have less mean residual values but because of their high minimum and maximum residual values, they have over forecasted. A

similar finding is seen with S curve except that it has under forecasted. Comparing the remaining two models (Linear and logarithmic models), it has been found that logarithmic model has high  $R^2$ , less p and residual values. Therefore, it is the best fitted and validated model.

Table 5 Model Validation

Model	Residual		
	Mean	Minimum	Maximum
Linear	8.21	-37.4	52.9
Logarithmic	5.32	-37.2	45.4
Inverse	2.17	-50.4	73.1
Quadratic	-0.81	-20.6	28.3
Cubic	-0.96	-20.4	28.7
Compound	-0.35	-127.6	68.3
Power	3.08	-66.54	56.4
S	-3.58	-343.8	81.8
Growth	-0.35	-127.6	68.3
Exponential	-0.35	-127.6	68.3

Figure 2 Year wise Estimated and Observed cases



**Discussion**

Palliative radiotherapy plays a vital role in all comprehensive palliative care programs. It has been estimated that approximately one half of all radiotherapy is prescribed with palliative intent. However, there is little information about the use of palliative radiotherapy in routine practice at population level<sup>12</sup>. In Manipal Teaching Hospital, 32.36% of all cancer cases received palliative care during the time period between 2000 and 2008. For the period of 2009-2015, in our study it is forecasted that 440 (28.22%) cancer patients would require palliative care out of the total forecasted 1559 cancer cases<sup>8</sup>. In our study, Table 1 shows that palliative radiotherapy can be used for numerous malignancies and in many clinical situations which is confirmed by published literature on the subject<sup>13-18</sup>. In our

study, we found that 37% of patients who received palliative care had carcinoma lung and among the entire number of lung cancer cases about 63.6% received palliative radiotherapy. Starting palliative care early in the treatment of patients with lung cancer can boost their quality of life and survival<sup>19</sup>.

**Modelling and Extrapolation:** A plot is a graphical representation of the collected data (independent and dependent variables) involved in a study. The association between these variables is then assessed by connecting the 'points' with a line. Though very accurate, this association cannot be relied upon to predict the future trend of this data. Now a 'model', which 'fits best' to the observed data has to be worked out. This is then 'fitted' and used to replace the existing set of data points as 'the appropriate model'. After 'modelling' the observed data, this model can be used to predict the future trends of the dependent variable for a given change in the other. The foregoing statement covertly mentions several requirements which often ensure confident achievement in any subsequent extrapolation from the model. The model selected must be the most appropriate one for the collected data. A usable and understandable curve-fitting method must be available from which the model facts those are reflective of future behaviour can be obtained<sup>20,21</sup>. Using such a curve fitting method, the number and trend of cancer cases receiving palliative treatment at Manipal Teaching Hospital from the year 2000 to 2015 were estimated. The statistical estimations have been done using different approaches in different situations. A regression equation based empirical approach can be applied when incidence rates and mortality data are available<sup>22</sup>. The decision regarding the selection of a suitable prediction approach is governed by the relative performance of the models for monitoring and prediction, and the phenomenon under study should be adequately interpreted<sup>23</sup>. In our study, the Logarithmic model was the best fitted model for observed data of cancer cases receiving palliative treatment (Fig 1). It may have over- or under- estimated the trends at some points, though not in all situations. A sudden annual decrease or increase in the trend is possible, as the curve cannot exactly reproduce these data points with its shape (Table 3). In this study, the future annual estimated cancer cases for palliative treatment (Table 3) shows an increasing trend after the year 2010. Such an increase might be likely as cancer incidence in developing countries is expected to rise principally due to the possible decline of mortality from infectious diseases, improved healthcare, population growth and increasing life expectancy<sup>2</sup>. The outcome of this study reiterates the applicability of statistical modelling in predicting the future trend of cancer patients seeking palliative care in the Nepalese context.

While selecting the model with the highest R-squared is tempting, it is not the recommended method. For instance, a cubic model will always have a higher R-squared than a quadratic model. The recommended method for selecting the best model is by cross-validation. That is, to compare

the predicted datasets obtained from each model with the original data set available with simultaneous comparison of the R-squares of the respective equations. Alternatively, the determination may be made graphically by overlaying sequence plots of both models for the observed dataset<sup>24,25,26</sup>. This paper has described a novel method of using the curve fitting method in number of cancer patients requiring palliative treatment. The approach is simple to understand and apply, and is capable of curve fitting a whole range of different models. It also has the advantage that several different models, for a given data series, can be easily investigated, thus easing the model selection dilemma.

### Conclusion

Most low-income and middle-income countries have health systems that are poorly prepared to grapple with the double burden of infectious and non-communicable diseases such as cancer. Poor economies cannot currently access curative therapies, state-of-the-art surgery, or expensive cancer drugs that are the mainstay of cancer care in developed nations. The projections of this hospital data show a continuously increasing trend in cancer patients receiving palliative care. It can be anticipated that the number of cancer patients requiring palliative treatment is on the rise in Nepal. Thus, in the near future, there might be a discrepancy between the requirement and the delivery of health care for these patients. This matter has to be given due consideration by the health authorities in order to provide adequate palliative care facilities in the country and enhance program planning and policy development.

### What this study adds

The outcome of this study reiterates the applicability of statistical modelling in predicting the future trend of cancer patients seeking palliative care.

### Conflict of Interests

The authors do not have any conflict of interest arising from the study.

### Acknowledgements

Dr. B M Nagpal, CEO Manipal Education and Medical group Dean, Manipal College of Medical Sciences, P O Box No 155, Deep Heights Pokhara (Nepal) for permitting the authors to use the hospital documents during the study.

---

### References

1. World Health Organization. Cancer Pain Relief and Palliative Care. (Technical report series 804). Geneva: World Health Organization; 1990.
2. Watson MS, Lucas CF, Hoy AM, Back IN. Oxford handbook of palliative care. 1<sup>st</sup> ed. Oxford : Oxford University Press, 2005, 85, 124.
3. Jassem J, Ramlau R, Santoro A, Schuette W, Chemaissani

- A, Hong S et al. Phase III trial of pemetrexed plus best supportive care compared with best supportive care in previously treated patients with advanced malignant pleural mesothelioma. *J Clin Oncol*. 2008;26(10):1698-704.
4. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. GLOBOCAN 2008: Cancer Incidence, Mortality Worldwide: IARC. CancerBase No. 10 [Internet]. Lyon, France: IARC; 2010. [Cited 10 Aug 2010] Available from URL: <http://globocon.iarc.fr>
  5. Current Hospice Care Patients, page 14. [online] February 2004 [cited 2010 June 15]. Available from: URL: <http://www.cdc.gov/nchs/data/nhhcsd/curhospicecare00.pdf>
  6. Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG, eds. *Cancer Registration: Principles and Methods*. Lyon: IARC Sci Publ No 95, 1991, Chapters 1.
  7. Pradhananga KK, Baral M, Shrestha BM. Multi-institution hospital-based cancer incidence data for Nepal - an initial report. *Asian Pac J Cancer Prev* 2009;10(2): 259-62.
  8. Sathian B, Bhatt CR, Jayadevan S, Ninan J, Baboo NS, Sandeep G. Prediction of cancer cases for a hospital in Nepal: a statistical modelling. *Asian Pac J Cancer Prev* 2010;11(2):441-5.
  9. Petrie A, Sabin C, *Medical Statistics at a glance*. 1<sup>st</sup> ed. London: Wiley-Blackwell Science, 2000, pp 71-4.
  10. Beyer WH. *Standard Mathematical Tables*. Cleveland: CRC Press, Inc., 1976.
  11. Chambers JM, Cleveland WS, Kleiner B, Tukey PA. *Graphical Methods for Data Analysis*. Boston, MA: Duxbury Press, 1983.
  12. Frodin JE, Jonsson E, Moller T, Werko L. Radiotherapy in Sweden—a study of present use in relation to the literature and an estimate of future trends. *Acta Oncol* 1996; 35(8): 967-79.
  13. Huang J, Zhou S, Groome P, Tyldesley S, Zhang-Solomans J, Mackillop WJ. Factors affecting the use of palliative radiotherapy in Ontario. *J Clin Oncol* 2001; 19(1):137–44.
  14. Johnston GM, Boyd CJ, Joseph P, MacIntyre M. Variation in delivery of palliative radiotherapy to persons dying of cancer in nova scotia, 1994 to 1998. *J Clin Oncol* 2001;19(14): 3323–32.
  15. Samant RS, Fitzgibbon E, Meng J, Graham ID. Family physicians' perspectives regarding palliative radiotherapy. *Radiother Oncol* 2006; 78(1): 101–6.
  16. Vulto JC, Louwman WJ, Poortmans PM, Coebergh JW. Hospital variation in referral for primary radiotherapy in South Netherlands, 1988–1999. *Eur J Cancer* 2005;41(17): 2722–7.
  17. Gaze MN, Kelly CG, Kerr GR, Cull A, Cowie VJ, Gregor A et al. Pain relief and quality of life following radiotherapy for bone metastases: a randomized trial of two fractionation schedules. *Radiother Oncol* 1997; 45:109–16.
  18. Awan AM, Weichselbaum RR. Palliative radiotherapy. *Hematol Oncol Clin North Am* 1990;4(6):1169–81.
  19. Temel JS, Greer JA, Muzikansky A, Gallagher ER, Admane S, Jackson VA et al. Early palliative care for patients with metastatic non-small-cell lung cancer. *N Engl J Med* 2010;363(8):733-42.
  20. Hindmarch GW, Towill DR. *Theory and Application of the Time Constant Learning Curve Model*. ORSA/TIMS, Puerto Rico, 1974.
  21. Nie NH, *SPSS: Statistical Package for the Social Sciences*. 2<sup>nd</sup> ed. New York: McGraw-Hill, 1975.
  22. Yang L. Estimating Cancer Burden in China. [online] 2005 [cited 2010 Feb 13]. Available from: URL:<http://acta.uta.fi/pdf/951-44-6475-3.pdf>
  23. Nobre FF, Monteiro AB, Telles PR, Williamson GD. Dynamic linear model and SARIMA: a comparison of their forecasting performance in epidemiology. *Stat Med* 2001;20(20):3051-69.
  24. Cuthbert D, Wood FS. *Fitting Equations to Data: Computer Analysis of Multifactor Data*. 2nd edition, New York, NY: Wiley-Interscience, 1999.
  25. Steiger JH, Browne MW. The comparison of interdependent correlations between optimal linear composites. *Psychometrika* 1984; 49: 11-24.
  26. Xue JL, Ma JZ, Louis TA, Collins AJ. Forecast of the number of patients with end-stage renal disease in the United States to the year 2010. *J Am Soc Nephrol* 2001;12(12): 2753–8.