

Estimation of Cancer Deaths in Korea for the Upcoming Years

Since the cancer has been the leading cause of deaths in Korea, estimation of the cancer deaths for the upcoming years in the population using the vital statistics is considered to be necessary. The aim of this study was to estimate the number and trends of cancer deaths in Korea. The expected numbers of cancer deaths were calculated by a time series model fitting the actual numbers of cancer deaths for each of the years 1983 through 2000 reported by Korea National Statistical Office. The options selected for the time series model included a quadratic time trend, which incorporated long-term information into the model and an autoregressive component which incorporated information about short-term fluctuations. The forecasting numbers of cancer deaths and their 95% confidence intervals were estimated for both genders and primary sites. The forecasting number of deaths from all cancers is increasing so that the cumulative number of expected cancer deaths between 2001 and 2005 would be about 309 thousand persons. Cancers of the lung, stomach, liver, and colorectum continue to be the most common causes of cancer deaths. While the numbers of expected cancer deaths in the stomach and liver show a decreasing trend, the cancer in the lung, colorectum, pancreas, breast, and oral cavity have an increasing trend. These observations indicate that cancer deaths in the near future would be increasing through the early 2000s, and there should be some urgent government's policy on the cancer management.

Key Words : Forecasting; Vital statistics; Death; Neoplasms; Korea

Jong-Myon Bae*, Kyu-Won Jung*,
Young-Joo Won*

Cancer Registration & Biostatistics Branch*,
National Cancer Center, Goyang; Department of
Preventive Medicine[†], Cheju National University
College of Medicine, Cheju, Korea

Received : 21 March 2002
Accepted : 23 May 2002

Address for correspondence

Jong Myon Bae, M.D.
Cancer Registration & Biostatistics Branch, National
Cancer Center, 809 Madu 1-dong, Ilsan-gu,
Koyang 411-351, Korea
Tel : +82.31-920-2131, Fax : +82.31-920-2139
E-mail : jmbae@ncc.re.kr

*This work was supported by Research Grant
N01E020 from the National Cancer Center, Korea.

INTRODUCTION

Since the cancer has been the leading cause of deaths in Korea since the 1980s (1), estimating the cancer deaths for the upcoming years is necessary for the nationwide cancer control, health resources planning, cancer surveillance, and quality control for cancer registries (2). The aim of this study was to estimate the cancer deaths for the upcoming years using the vital statistics.

MATERIALS AND METHODS

The primary source of information about cancer deaths in Korea was the annual report on the cause of death statistics based on vital registration, which was compiled and published by Korea National Statistical Office (KNSO) (3). To estimate the number of cancer deaths expected to occur during the upcoming years, the mathematical modeling involved the actual number of cancer deaths reported by KNSO for each of the years 1983 through the most recent available year. The notation used for cancer deaths involves 'A' for actual counts of

cancer deaths, 'F' for forecasted counts of cancer deaths, and 'yy' for the year. Thus 'A_99' refers to the actual number of cancer deaths in 1999 and 'F_05' refers to the forecasted counts of deaths in 2005.

We estimated the number of cancer deaths expected to occur in Korea for upcoming years using a two-step procedure (4). First, the actual number of cancer deaths (A_83, A_84, ..., A_00) for each of the years 1983 through 2000 were fitted to a time series model in the SAS procedure PROC FORECAST (5) (Appendix). The model was then used to forecast an estimate of the number of cancer deaths expected to occur in Korea for the upcoming years (F_01, F_02, ..., F_10) and to calculate the 95% confidence limits for the estimates. The options selected for the time series model included a linear or quadratic time trend, which incorporated long-term information into the model and an autoregressive component, which incorporated information about short-term fluctuations. To specify the degree of the time trend model was based on a comparison of the R-square statistics of observation (6) (Table 1). Estimates of the number of cancer deaths in primary sites was calculated using the same modeling process that was used for all cancer deaths.

RESULTS

Annual observed and expected cancer deaths are given in Fig. 1. Table 2 summarizes the expected number of cancer deaths as well as deaths in the ten major sites of cancer-lung, stomach, liver, colorectum, pancreas, esophagus, leukemia, breast, head and neck, and brain for both genders combined in Koreans. The forecasting number of cancer deaths might be increasing so that the cumulative number of expected cancer deaths between 2001 and 2005 would be 309,839 (95% CI: 286,039-333,639). Cancers of the lung, stomach, liver, and colorectum continue to be the most common causes of cancer deaths. These four cancers account for more than 60% of the total cancer deaths. Especially, the lung cancer would continuously be the leading cause of cancer death in the near

Table 1. Root mean square errors for specifying the degree of the time trend model

Model	Root mean square error
Quadratic trend with autoregression	1261.2
Linear trend	1420.8
Quadratic trend	1403.8
Linear exponential smoothing	1478.1
Double exponential smoothing	1808.6
Damped Trend exponential smoothing	1476.2
Random walk with drift	1515.4
Log linear trend	1916.4
Log linear exponential smoothing	1547.1
Log double exponential smoothing	1869.3
Log random walk with drift	1597.2

future. Through 2005, the expected number of deaths in cancer of the lung, colorectum, pancreas, breast, and oral cavity would be increasing. But cancer deaths in the stomach, liver, leukemia, and brain would show a decreasing trend (Fig. 2).

Table 3 shows the expected number of all cancer deaths as well as deaths in the six major sites of cancer-lung, stomach, liver, colorectum, pancreas, and prostate in Korean men. The forecasting number of cancer deaths might be increasing through 2005. Among the six major cancers in men, the lung cancer would continuously be the leading cause of cancer

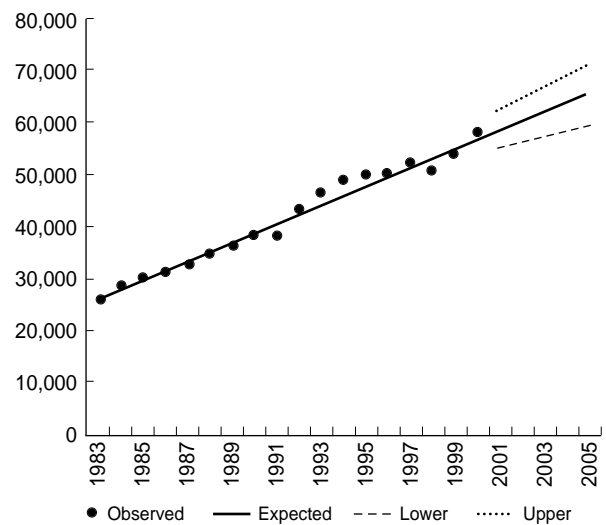


Fig. 1. Annual observed and expected cancer deaths and those 95% confidence intervals.

Table 2. Expected numbers of cancer deaths (and 95% confidence intervals) in both genders in Korea

Site (ICD-10)	2001	2002	2003	2004	2005	Sum (%)*
All cancers (C00-C97)	58,678 (54,830 - 62,526)	60,341 (56,127 - 64,556)	61,986 (57,315 - 66,657)	63,613 (58,395 - 68,830)	65,221 (59,372 - 71,070)	309,839 (100.0)
Lung (C34)	12,097 (11,318 - 12,876)	12,791 (11,938 - 13,644)	13,499 (12,553 - 14,444)	14,220 (13,164 - 15,276)	14,954 (13,770 - 16,138)	67,560 (21.8)
Stomach (C16)	11,507 (10,357 - 12,658)	11,455 (10,084 - 12,825)	11,392 (9,841 - 12,944)	11,321 (9,578 - 13,064)	11,241 (9,284 - 13,198)	56,916 (18.4)
Liver (C22)	9,761 (9,067 - 10,456)	9,646 (8,885 - 10,406)	9,494 (8,651 - 10,337)	9,305 (8,364 - 10,247)	9,081 (8,025 - 10,137)	47,287 (15.3)
Colorectum (C18-C21)	4,440 (4,202 - 4,678)	4,802 (4,542 - 5,063)	5,180 (4,891 - 5,469)	5,574 (5,252 - 5,897)	5,984 (5,623 - 6,364)	25,982 (8.4)
Pancreas (C25)	2,870 (2,674 - 3,066)	3,043 (2,829 - 3,258)	3,220 (2,983 - 3,458)	3,401 (3,135 - 3,666)	3,584 (3,287 - 3,882)	16,119 (5.2)
Esophagus (C15)	1,512 (1,382 - 1,642)	1,523 (1,381 - 1,666)	1,531 (1,373 - 1,688)	1,534 (1,357 - 1,710)	1,533 (1,335 - 1,730)	7,633 (2.5)
Leukemia (C91-C95)	1,342 (1,169 - 1,516)	1,330 (1,140 - 1,521)	1,314 (1,103 - 1,525)	1,294 (1,059 - 1,530)	1,270 (1,006 - 1,534)	6,551 (2.1)
Breast (C50)	1,258 (1,157 - 1,359)	1,328 (1,218 - 1,438)	1,400 (1,278 - 1,522)	1,474 (1,338 - 1,611)	1,551 (1,398 - 1,704)	7,011 (2.3)
Head & Neck (C00-C14)	830 (567 - 1,093)	900 (613 - 1,188)	974 (655 - 1,293)	1,052 (696 - 1,408)	1,133 (734 - 1,533)	4,890 (1.6)
Brain (C70-C72)	853 (414 - 1,292)	721 (230 - 1,213)	719 (132 - 1,305)	700 (41 - 1,359)	588 (0 - 1,338)	3,581 (1.2)

*Sum (%): the cumulated numbers of cancer deaths between 2001 and 2005 and its column percent.

deaths. The trends of cancer deaths by primary site would be as same as those in both genders (Fig 3).

Table 4 shows the expected number of all cancer deaths as well as deaths in the seven major sites of cancer-lung, stomach, liver, colorectum, pancreas, breast, and uterine cervix in Korean women. The forecasting number of cancer deaths might be increasing through 2005 too. Although the stomach cancer would have a decreasing trend in death numbers, it would be the first leading cause of cancer deaths through 2005 (Fig. 4). The primary sites having increasing trends in cancer deaths would be the lung, colorectum, pancreas, breast, and uterine cervix.

DISCUSSION

Measurement of the burden of cancer in a population is essential in order to decrease the deaths from cancer. The Ameri-

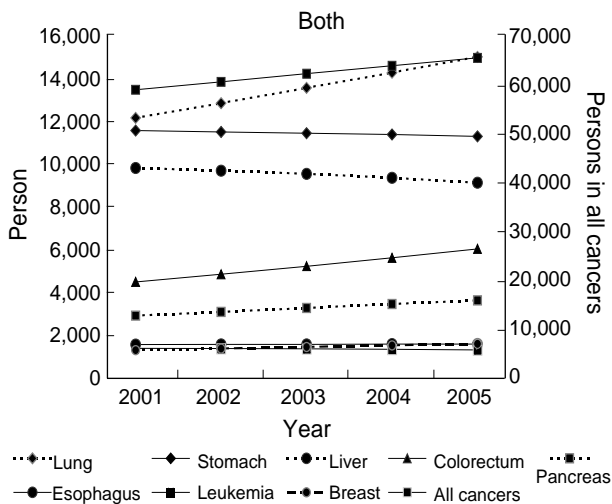


Fig. 2. Trends of expected cancer deaths in both in Korea.

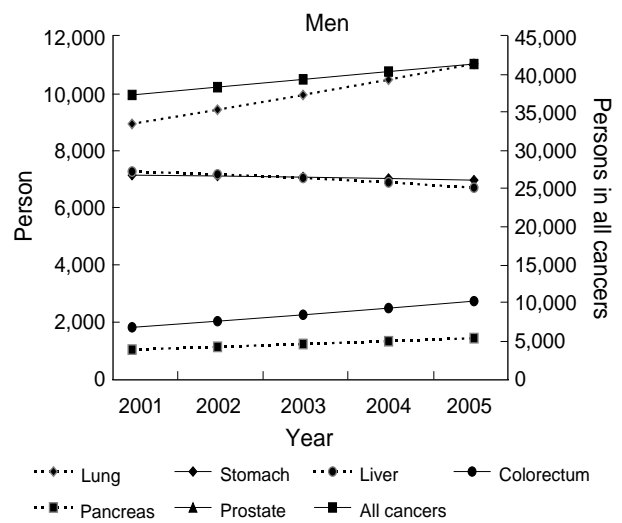


Fig. 3. Trends of expected cancer deaths in Korean men.

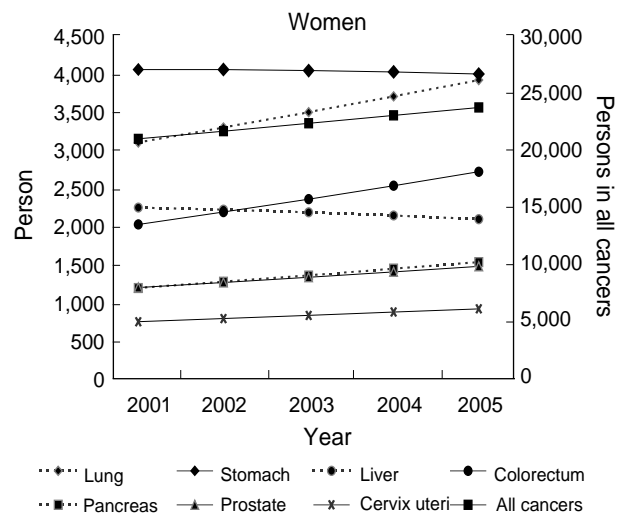


Fig. 4. Trends of expected cancer deaths in Korean women.

Table 3. Expected numbers of cancer deaths (and 95% confidence intervals) in Korean men

Site (ICD-10)	2001	2002	2003	2004	2005	Sum (%)*
All cancers (C00-C97)	37,403 (35,100 - 39,706)	38,383 (35,861 - 40,906)	39,342 (36,546 - 42,138)	40,278 (37,155 - 43,401)	41,193 (37,692 - 44,694)	196,599 (100.0)
Lung (C34)	9,033 (8,520 - 9,547)	9,492 (8,888 - 10,095)	9,977 (9,299 - 10,655)	10,480 (9,720 - 11,239)	10,995 (10,143 - 11,847)	49,977 (25.4)
Stomach (C16)	7,356 (6,678 - 8,035)	7,324 (6,581 - 8,068)	7,286 (6,462 - 8,110)	7,242 (6,322 - 8,162)	7,192 (6,160 - 8,223)	36,400 (18.5)
Liver (C22)	7,472 (6,937 - 8,007)	7,382 (6,796 - 7,968)	7,264 (6,615 - 7,914)	7,119 (6,393 - 7,844)	6,945 (6,132 - 7,758)	36,182 (18.4)
Colorectum (C18-C21)	2,371 (2,212 - 2,530)	2,570 (2,396 - 2,744)	2,778 (2,586 - 2,971)	2,996 (2,781 - 3,211)	3,223 (2,982 - 3,464)	13,938 (7.1)
Pancreas (C25)	1,638 (1,507 - 1,769)	1,729 (1,585 - 1,873)	1,821 (1,662 - 1,981)	1,915 (1,737 - 2,093)	2,010 (1,811 - 2,209)	9,113 (4.6)
Prostate (C61)	573 (523 - 622)	640 (586 - 694)	712 (651 - 772)	787 (720 - 855)	867 (791 - 942)	3,579 (1.8)

*Sum (%): the cumulated numbers of cancer deaths between 2001 and 2005 and its column percent.

Table 4. Expected numbers of cancer deaths (and 95% confidence intervals) in Korean women

Site (ICD-10)	2001	2002	2003	2004	2005	Sum (%)*
All cancers (C00-C97)	21,275 (19,698 - 22,853)	21,958 (20,230 - 23,686)	22,644 (20,729 - 24,560)	23,334 (21,195 - 25,474)	24,028 (21,630 - 26,426)	113,239 (100.0)
Lung (C34)	3,143 (2,892 - 3,394)	3,340 (3,065 - 3,615)	3,543 (3,238 - 3,847)	3,751 (3,410 - 4,091)	3,964 (3,583 - 4,346)	17,741 (15.7)
Stomach (C16)	4,101 (3,549 - 4,653)	4,102 (3,443 - 4,760)	4,090 (3,343 - 4,836)	4,070 (3,231 - 4,908)	4,044 (3,102 - 4,985)	20,407 (18.0)
Liver (C22)	2,289 (2,109 - 2,469)	2,263 (2,066 - 2,461)	2,229 (2,010 - 2,448)	2,187 (1,943 - 2,431)	2,136 (1,862 - 2,410)	11,104 (9.8)
Colorectum (C18-C21)	2,069 (1,948 - 2,190)	2,232 (2,100 - 2,365)	2,402 (2,255 - 2,549)	2,579 (2,415 - 2,743)	2,762 (2,578 - 2,946)	12,044 (10.6)
Pancreas (C25)	1,232 (1,144 - 1,320)	1,314 (1,218 - 1,411)	1,399 (1,292 - 1,506)	1,486 (1,366 - 1,605)	1,574 (1,440 - 1,709)	7,005 (6.2)
Breast (C50)	1,237 (1,134 - 1,339)	1,304 (1,192 - 1,416)	1,374 (1,250 - 1,498)	1,445 (1,307 - 1,584)	1,519 (1,364 - 1,674)	6,879 (6.1)
Cervix uteri (C53)	788 (693 - 883)	830 (726 - 934)	873 (757 - 988)	915 (787 - 1,044)	959 (814 - 1,103)	4,365 (3.9)

*Sum (%): the cumulated numbers of cancer deaths between 2001 and 2005 and its column percent.

can Cancer Society (ACS) has estimated the number of new cancer cases and cancer deaths that are expected to occur in the whole United States and individual states during the upcoming years (7, 8). Using the current methods for estimation of the ACS (4), we tried to estimate the number of cancer deaths that are expected to occur in Korea in the near future.

However, the estimated number of cancer deaths for the upcoming years has limitations and should be interpreted with caution when used to study trends in cancer mortality (4). Although the estimates are based on the most recent available cancer mortality data, the estimated numbers of deaths are computed before the year begins so that these estimates may vary considerably from year to year. On the other hand, unanticipated changes to cancer survivals (for example, activation of cancer screening programs, development of a new modality of cancer diagnosis or treatments, etc), could be a factor for instability of these expectations.

The problem of choosing a forecast approach depends on the relative performance of the models for monitoring and prediction, with an adequate interpretation of the phenomenon under study (9). Since selection of the models could be based on the logarithm of the predictive likelihood error (10), this study used the R-square statistic and sigma types of observation in order to specify the degree of the time trend model.

Five possible choices for an estimate were considered as acceptable: the forecast itself, the upper or lower 95% confidence limit, and the midpoints between the forecast and the upper or lower confidence limit (3). When appropriate, the forecast itself with confidence limits was selected as the estimate to compensate for the effects of recently changing mortality rates or large year-to-year variations in the number of cancer deaths (11).

REFERENCES

1. National Statistical Office (Republic of Korea). *Annual report on the cause of death statistics (based on vital registration)*. Seoul: The Office; 2001.
2. Pickle LW, Feuer EJ, Edwards BK, National Cancer Institute. *Prediction of incident cancer cases in non-SEER counties. Proceedings of the Biometrics Section of the 2000 Annual Meeting of the American Statistical Association*. 2001: 45-52.
3. National Statistical Office (Republic of Korea). *Korea Statistical Information System*. Available from: URL: <http://kosis.nso.go.kr>.
4. Wingo PA, Landis S, Parker S, Bolden S, Health CW. *Using cancer registry and vital statistics data to estimate the number of new cancer cases and deaths in the United States for the upcoming year*. *J Reg Management* 1998; 25: 43-51.
5. SAS Institute Inc. *SAS/ETS User's Guide. Version 8*. Cary NC: SAS Institute Inc. 1999: 577-631.
6. SAS e-Intelligence. *Time series*. Available from: URL: <http://www.sas.com/korea>.
7. Greenlee RT, Hill-Harmon MB, Murray T, Thun M. *Cancer statistics, 2001*. *CA Cancer J Clin* 2001; 51: 15-36.
8. Jemal A, Thomas A, Murray T, Thun M. *Cancer statistics, 2002*. *CA Cancer J Clin* 2002; 52: 23-47.
9. Nobre FF, Monteiro ABS, Telles PR, Williamson GD. *Dynamic linear model and SARIMA: a comparison of their forecasting performance in epidemiology*. *Statist Med* 2001; 20: 3051-69.
10. Pole A, West M, Hamson J. *Applied Bayesian forecasting and time series analysis*. *Texts in Statistical Science*. Chapman & Hall: New York: 1994.
11. Wingo PA, Ries LAG, Rosenberg HM, Miller DS, Edwards BK. *Cancer incidence and mortality 1973-1995: a report card for the US*. *Cancer* 1998; 82: 1197-207.

Appendix

The FORECAST procedure provides a quick and automatic way to generate forecasts for many time series in one step. We adapt the stepwise autoregressive method. This method combines time trend regression with an autoregressive model and uses a stepwise method to select the lags to use for the autoregressive process as below,

$$\text{Death}_t = \mu + \alpha \cdot t + \beta \cdot t^2 + \nu_t, \text{ for } t=1983, 1984, \dots, 2000$$

where $\nu_t = \alpha_1 \cdot \nu_{t-1} + \alpha_2 \cdot \nu_{t-2} + \dots + \alpha_p \cdot \nu_{t-p} + \varepsilon_t$