Mortality and Approximate Hazard Plots[1]

Morteza Aalabaf-Sabaghi[2]

E.C.O. College of Insurance, Allameh Tabatabai University, No.1, Malek-o-Shoara Bahar, Taleghani Avenue, Tehran 15717, Iran

February 2009

Abstract

Hazard rates are traditionally estimated using actuarial methods from mortality tables. Here we estimate hazard rates from tables of death frequencies for grouped ages using non-parametric methods and compare them with parametric methods. We use proportional hazard function to take into account lags in mortality and duration dependence of ailments that take effect in time.

JEL: C14, J11

Keywords: Hazard rate, proportional hazard, duration dependence, survival function.

---

1. Introduction

Hazard rates are traditionally estimated using actuarial methods from mortality tables. Here we estimate hazard rates from tables of death frequencies for grouped ages using non-parametric methods and compare them with parametric methods. The hazard rate is a conditional probability that measures the instantaneous rate of failure. Using results in Aalabaf-Sabaghi (2002) we calculate a series of tabular hazard rates for major conditions that pose a health risk to individuals. On the other hand, in order to find a parametric hazard rate, we use the methodology of Corro (2002) that uses life expectancy. Life expectancy is shown to inversely relate to hazard rate for a particular parametric description of the survival distribution function. In a restrictive form, we adopt this methodology to find a first parametric estimate for hazard rates for major conditions using Iranian life tables. In this "approximation" life expectancy that represents mean time to failure is used to find hazard rates. It is understood in this approach that a restrictive parameterization of the survival distribution function, namely the constant hazard, can be shown to inversely relate to the life expectancy. However, we also take note that life expectancy reflects an expectation of life over a longer time horizon (at t=0) whereas the hazard rate is an instantaneous rate of death conditional on survival to that time.

In the following sections we outline our approach and then present graphical results of hazards to life for mortality causes comparing non-parametric and parametric results.

2. The model

In most applications, the hazard rate function is typically written as:

$$h(t) = \frac{f(t)}{1 - F(t)} \tag{1}$$

Where $F(t) = \int_0^t f(t)\, dt$ is the probability distribution of death hazard with density function $f(t)$ defined over time t and $S(t) = 1 - F(t)$ is the survival function. The shape of the hazard function in (1) depends on the particular probability distribution F (t) that best describes death probability from data available on mortality statistics. A number of probability distributions may be used; in particular the constant hazard $h(t) = h$, exponential function and log logistic are found in the literature. The choices of these distributions depend usually on the goodness-of-fit criteria. However, the use of proportional hazard function (Cox (1972)) allows in addition to incorporating individual heterogeneity factors to define a density function that varies over time. It may be desirable to define densities that have positive or negative time dependence. In Figure 1, we show the life expectancy for men in Iran. The hazard rates derived from these life expectancies are shown in Figure 2. The hazard rate declines until age 12 and then rises thereafter. It shows negative time dependence till age 12 and then exhibits positive time dependence as it can be seen in Figures 2 and 4. The proportional hazard model is quite adaptable to take into account time dependence. In particular it enables to define a density function such that hazard of death declines till age 12 and then rises thereafter. The

choice of age 12 as a cut-off point for our sample is based on observations in our data-set.

Suppose for a cohort of N_0 born at time t=0, a discrete random variable S_t representing the number of survivors at the end of time period t can be written as

$$S_t = N_0 \prod_{t=0}^{T}(1-\gamma_t) \qquad (2)$$

Where γ_t is the probability of death for period t.

Now we can treat (2) as a survival function as in Le (1997) and London (1997) and write a hazard function in terms of this survival function

$$h_t = \frac{\left(-\dfrac{d}{dt}S_t\right)}{S_t} \qquad (3)$$

Where h_t is the hazard faced by an individual at time t and (3) is the hazard function. But we can write (3) as

$$h_t = -\frac{d}{dt}\ln S_t \qquad (4)$$

Using (1) and substituting in (4) we can see that

$$h_t = -\sum_{t=0}^{T-1}\frac{d}{dt}\ln(1-\gamma_t) \qquad (5)$$

This derivation of the hazard rate has similarities to the derivation in Corro (2002) for inverting life expectancy.

In Corro (2002) for a particular parameterization it is shown that:

$$1+\frac{d\rho}{dt} = h(t)\rho(t) \qquad (6)$$

and therefore,

$$h(t) = \frac{1}{\rho(t)} + \frac{d(\ln \rho)}{dt} \qquad (7)$$

We use equation (7) and calculate hazard rates shown in Figure 2. In Figure 1 life expectancy is shown using life tables for Iran calculated in Zanjani and Noorallahi (1999). This is a typical life expectancy showing average number of years until death. The corresponding inverted life expectancy representing the hazard rate for the same population is shown in Figure 2 which is also a typical "bath tub" shape. In Figure 3 we have the mean hazard rate for our sample data. This data groups all deaths in 12 age groups.

In Figures 5 through 16 hazard rates for all age groups facing 18 major illnesses are shown. Figure 17 plots all age groups in the same diagram. It is very apparent from a visual inspection of all these figures that three major causes have the highest fatalities among the population. These are indicated by hazard rates showing peaks at cancer, diseases of the circulatory system (heart attack) and accidents and poisoning. The death category labeled as "unspecified" counts a small proportion of all deaths in all age groups except for the 0-1 age group. For the "age unknown" however, "unspecified" category is the highest among all causes of

deaths, although we still can visually see a marked difference for cancer, diseases of the circulatory system and accidents and poisoning. Appendix A, lists the 18 categories of major causes of death and the corresponding aggregation of an extensive list of 151 mortality causes that are tabled in Appendix B.

In Figures 18 through 35 we show hazard rates for all age groups. In all these figures the unknown age category may be ignored as this "unknown" age group may be presented in any section of the age axis. Again, the proportion of fatalities due in the "unknown" age group is small except for the "unspecified" causes. We also observe the usual "bath tub" shape for most of the hazard rates except for "mental disorders", "complications of pregnancy and child birth" and "accidents and poisoning". For these mortality causes the exponential decay is somewhat qualified. Using this tabular calculation of hazard has this advantage of observing differences in death rates due to different causes. While a parametric hazard rate needs to accommodate these differences, once we have obtained results of non-parametric estimation of the hazard rates. In Figure 36, we present hazard rates for all major causes that are proportionately drawn in the same scale, thus their differences are less pronounced and the usual "bath tub" shape re-emerges.

In Figures 37, 38 and 39 we show the disaggregated plots of hazard rates for conditions that are categorized under 'cancer", "diseases of the circulatory system" and "accidents and poisoning". Here proportionately fewer younger persons die from cancer or hart attacks. But accidents are much more frequent for the 25-34 year olds and the very old.

3. Conclusion

This short note shows how approximate hazard rates may be obtained from a set of life expectancy data. As hazard rate reflect local failure rates it may be approximated in this way for grouped data and for particular death frequencies. However, the limitation of this method needs to be taken into account in any application to data.
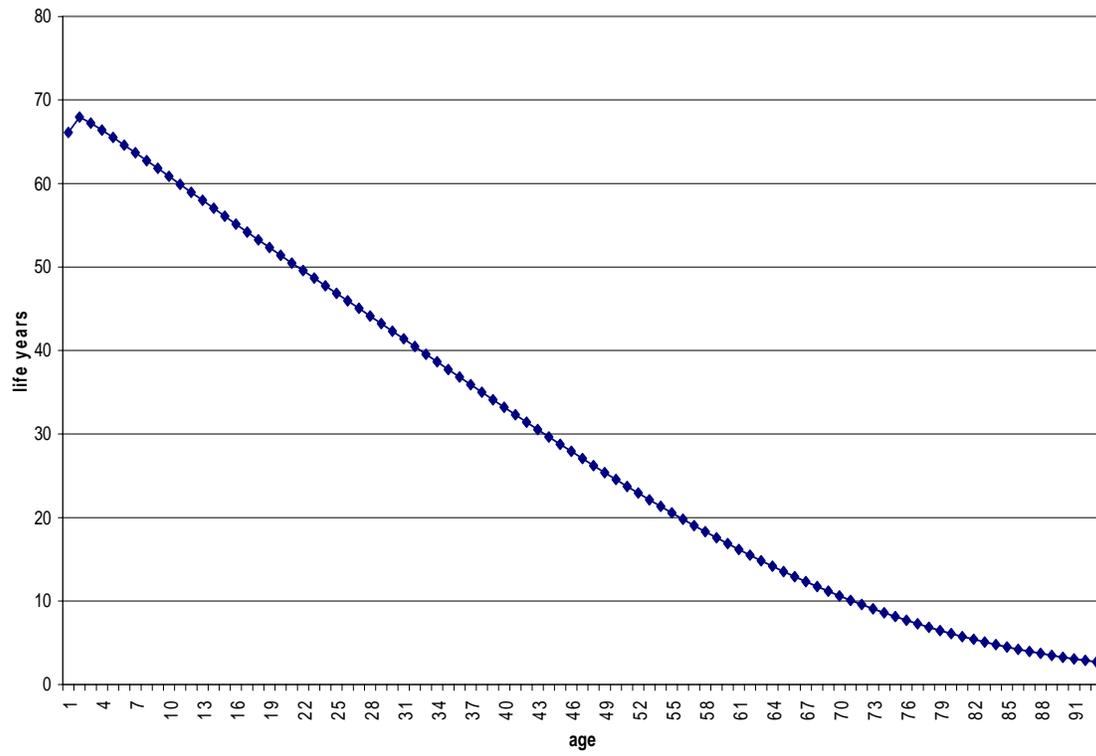
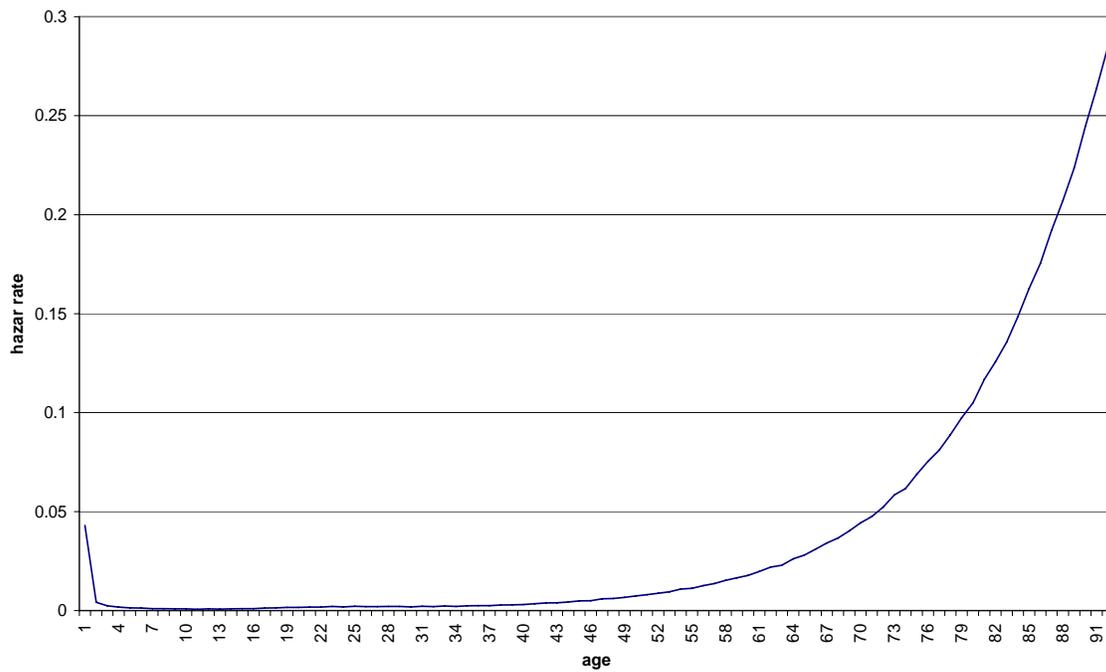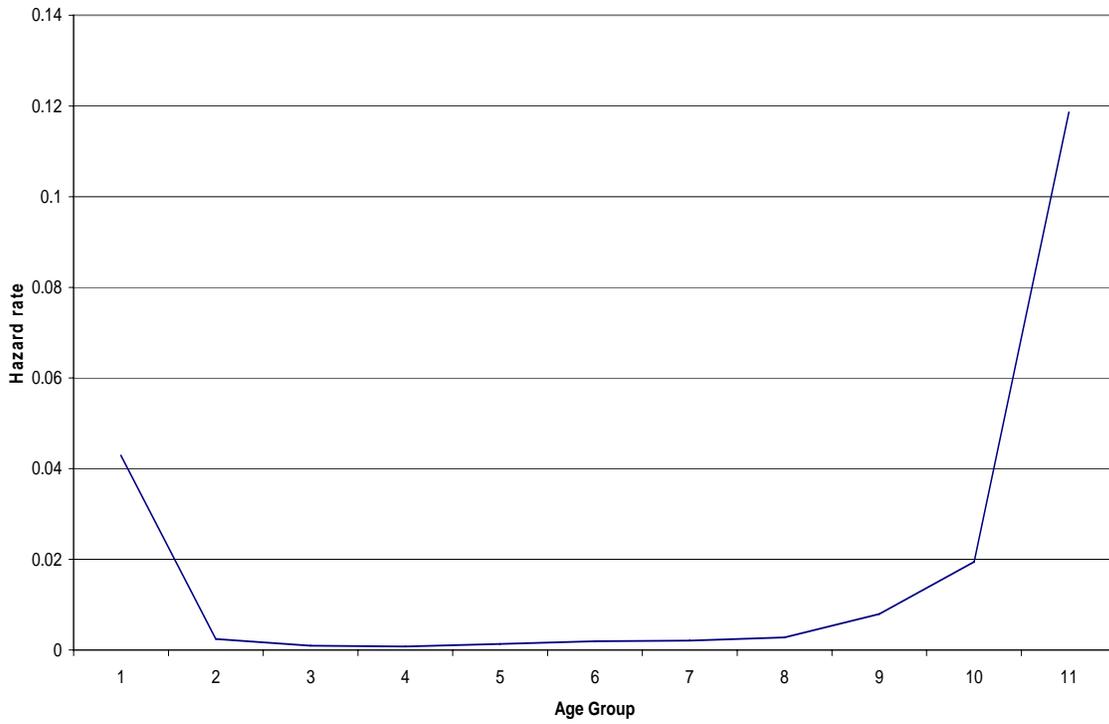**Figure 1. life expectancy**

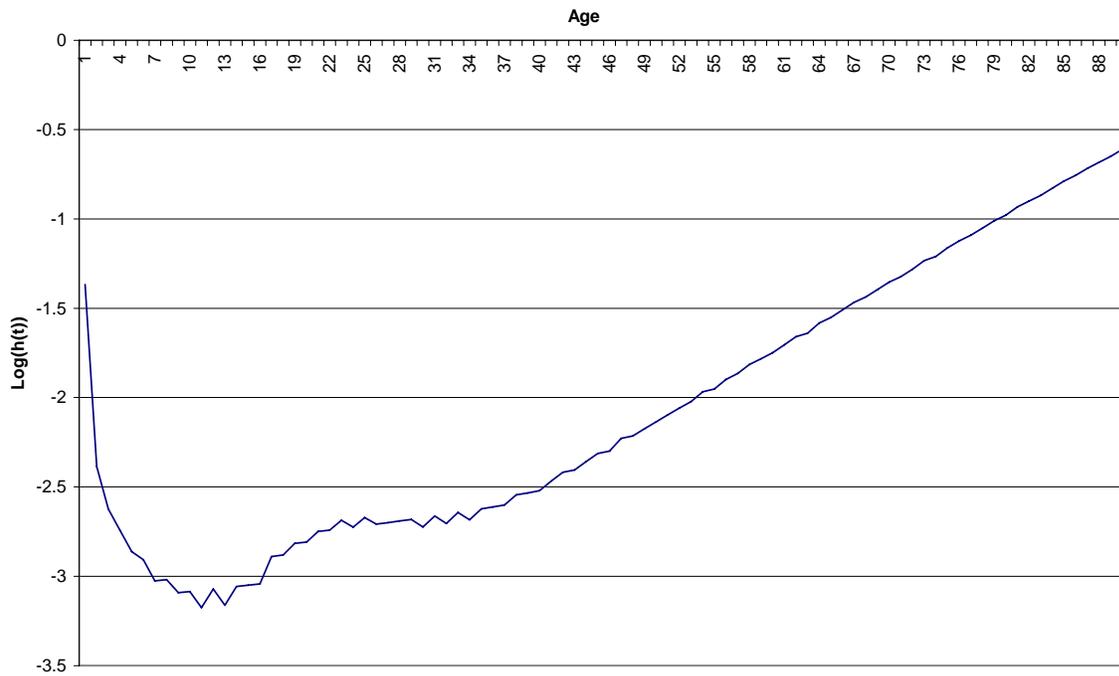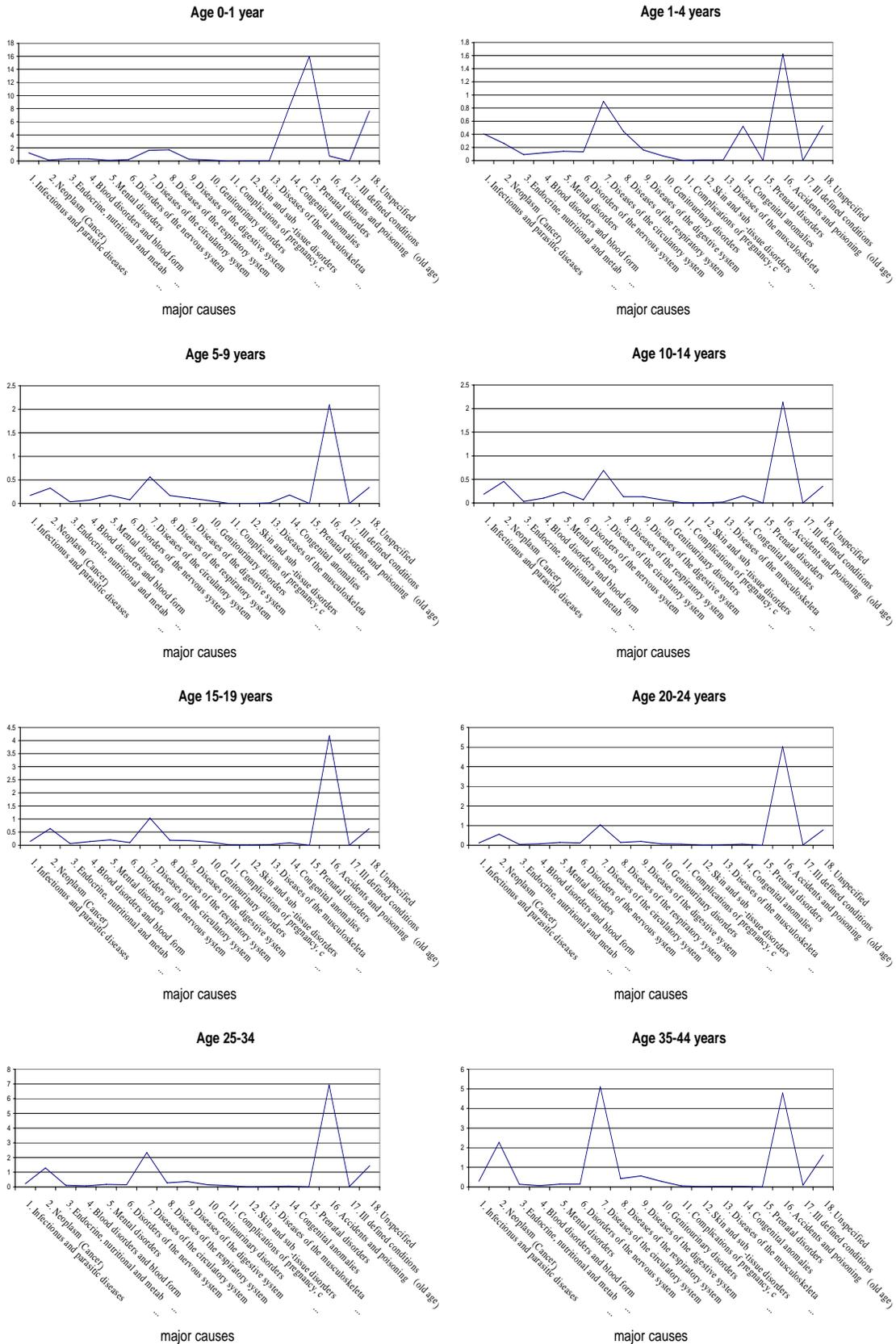**Figure 2. Hazar rate**

Figure 3. Mean Hazard Rate

Figure 4. Log(hazard rate)
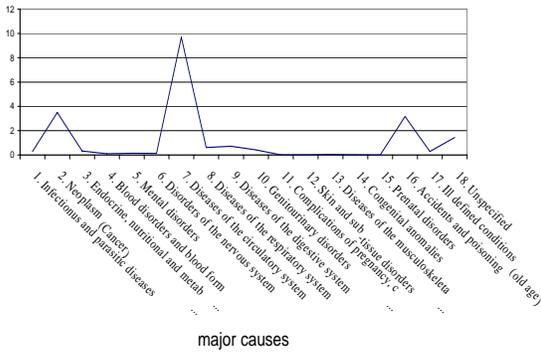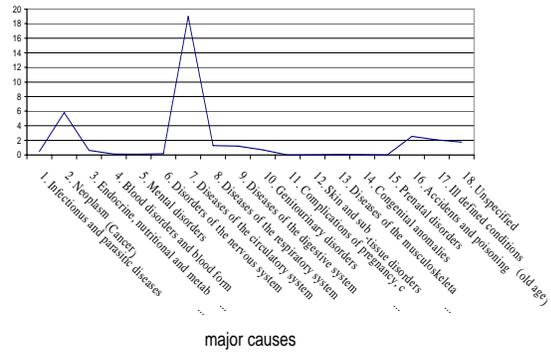

6

Fig 1: Hazard rate in different age groups

Age 0-1 year

major causes

Age 1-4 years

major causes

Age 5-9 years

major causes

Age 10-14 years

major causes

Age 15-19 years

major causes

Age 20-24 years

major causes

Age 25-34

major causes

Age 35-44 years

major causes

Age 45-54 years

major causes

Age 55-64 years

major causes

Age 65+
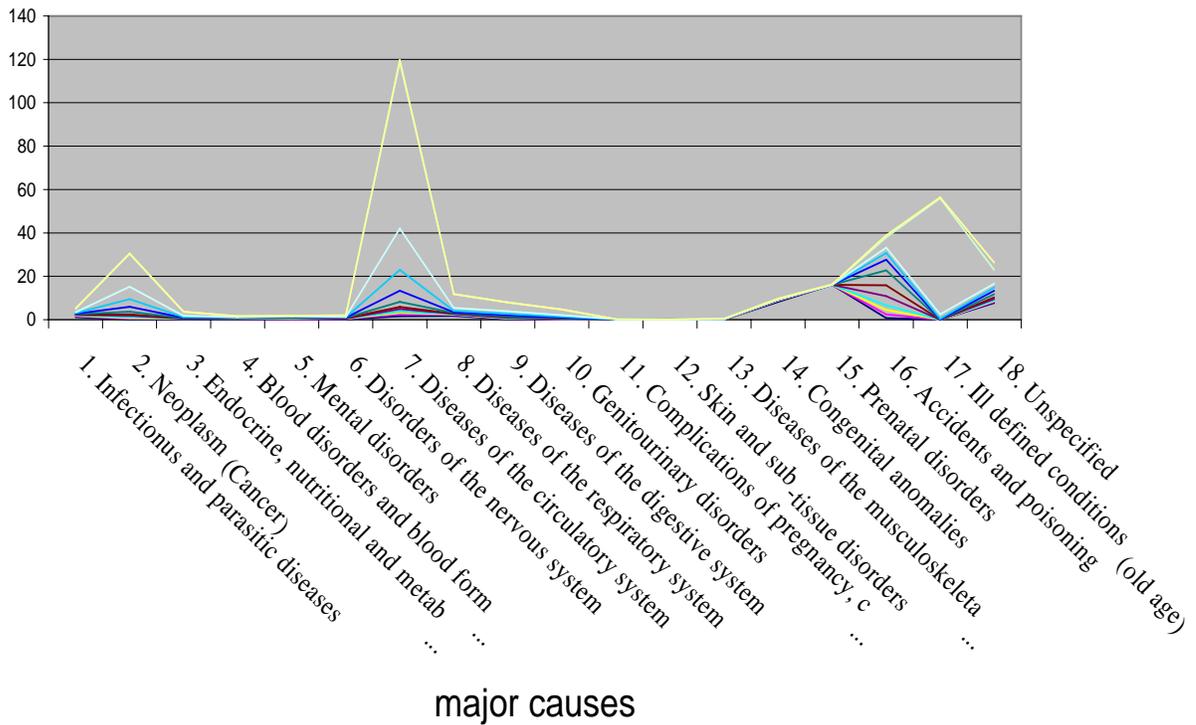

major causes

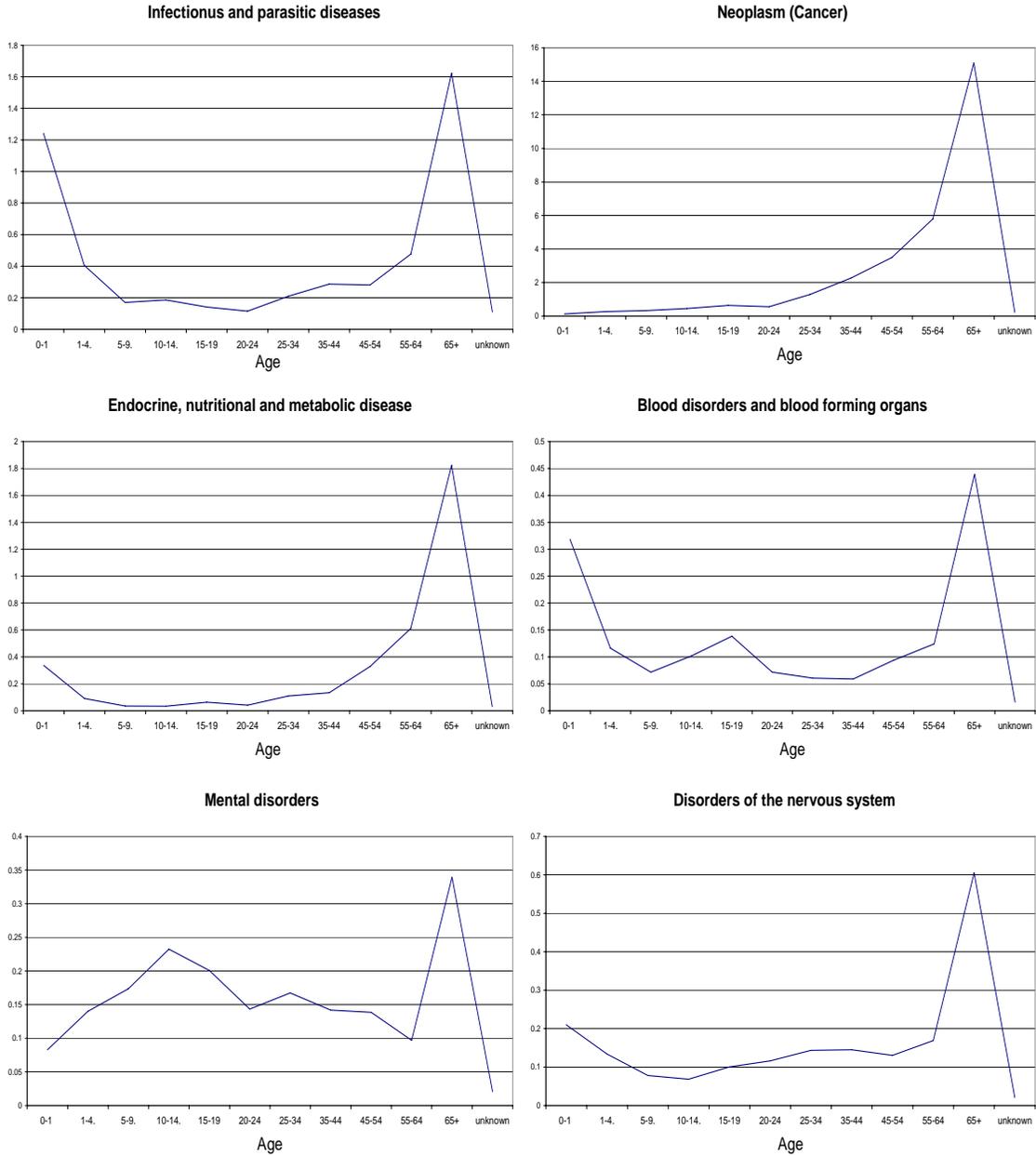**Age unknown**

major causes:
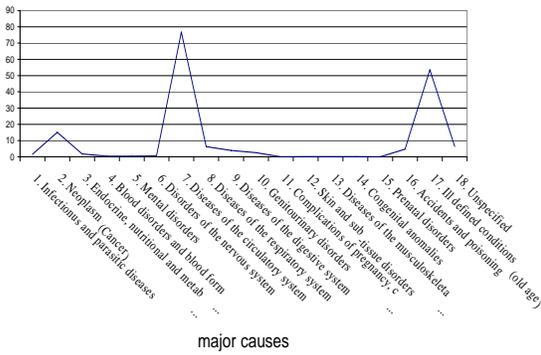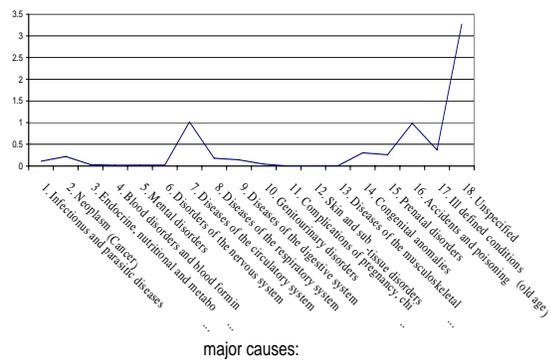
Hazard rates: all ages
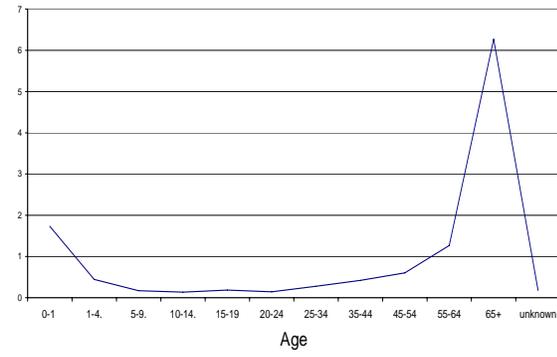

major causes

Fig 2: Hazard rate for different causes during a life-time

Infectionus and parasitic diseases

Neoplasm (Cancer)

Endocrine, nutritional and metabolic disease

Blood disorders and blood forming organs

Mental disorders
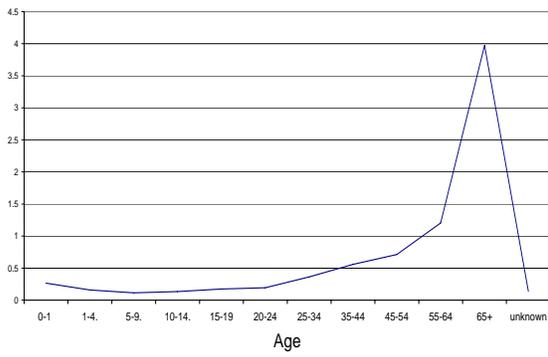

Disorders of the nervous system

Diseases of the circulatory system
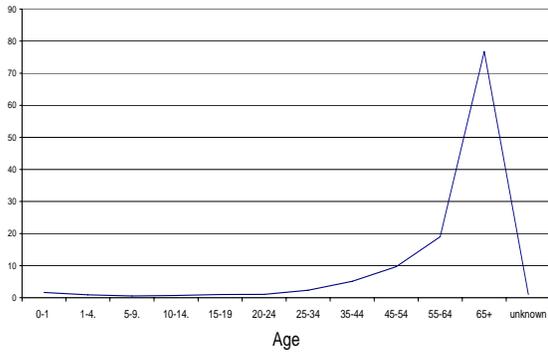
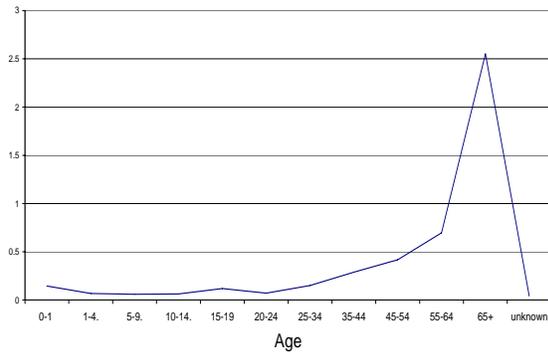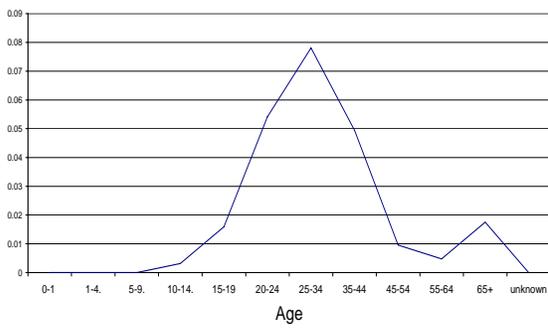**Diseases of the respiratory system**

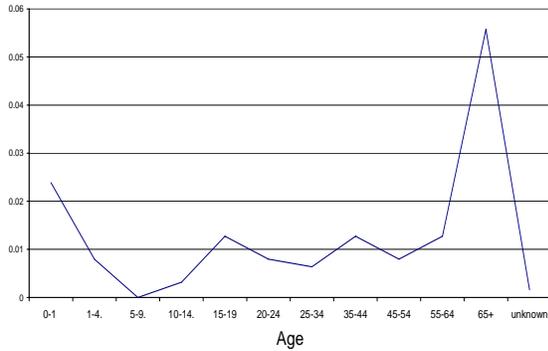**Diseases of the digestive system**
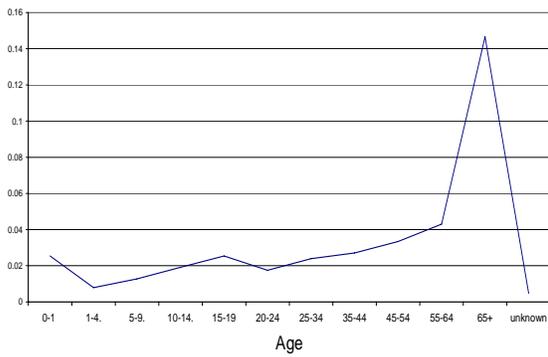

Genitourinary disorders

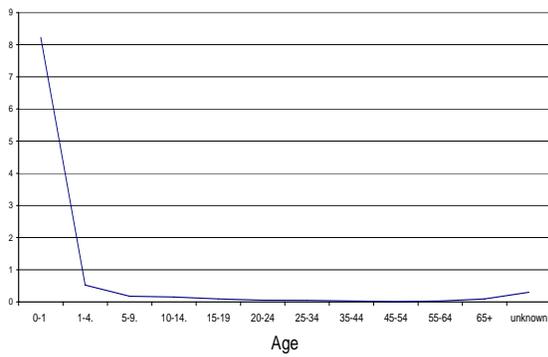**Complications of pregnancy, child birth and the puerperium**
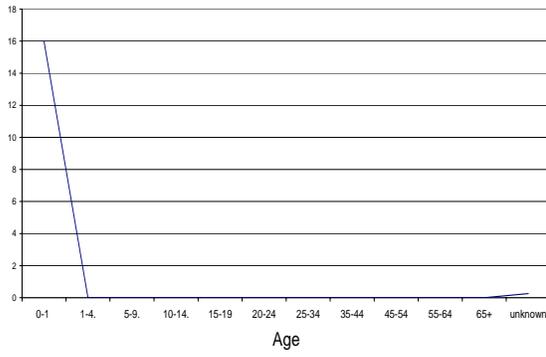

Skin and sub-tissue disorders
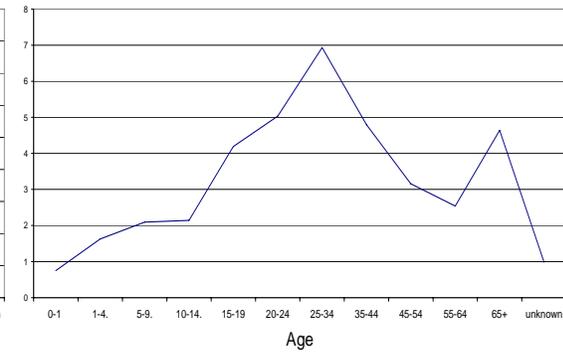

Diseases of the musculoskeletal system
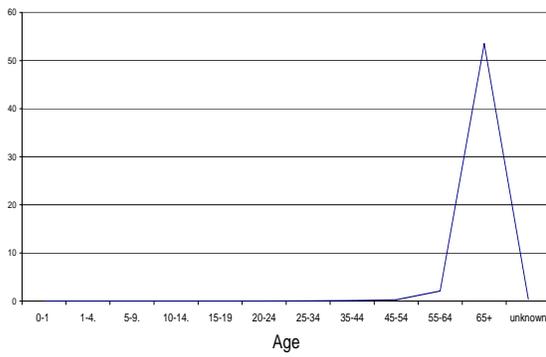

Congenital anomalies
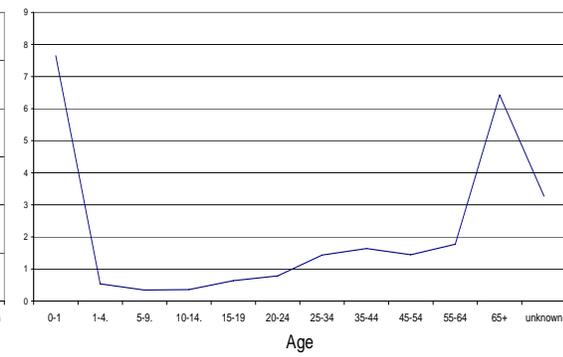

10

Prenatal disorders
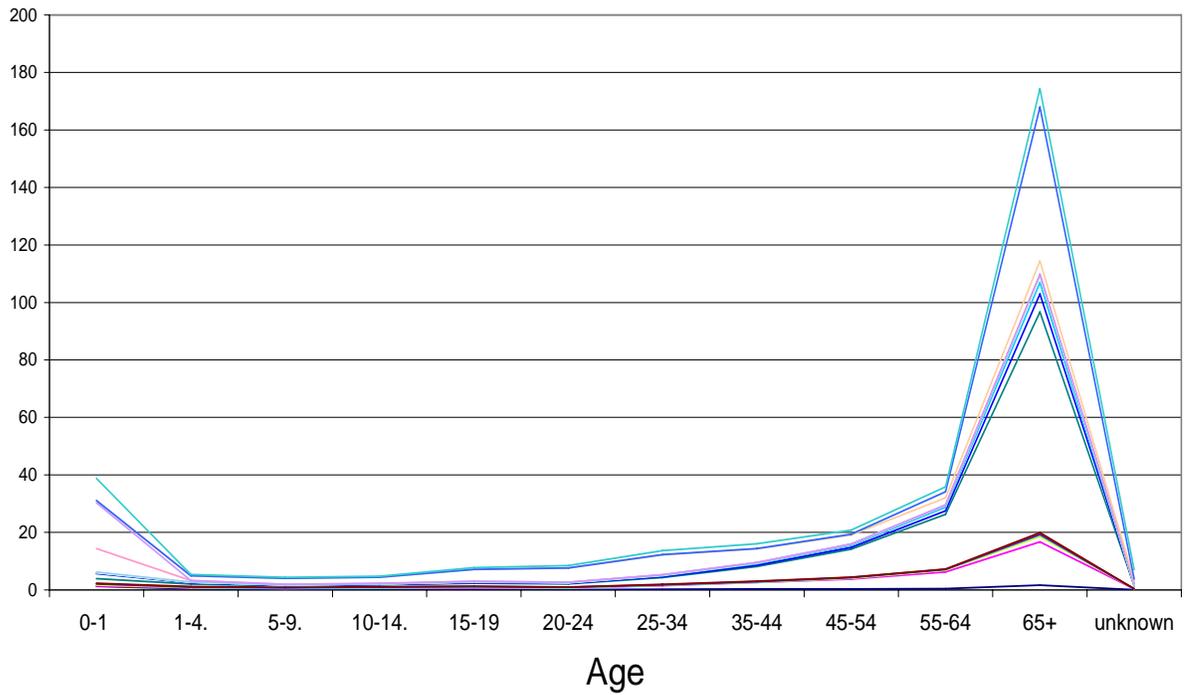
Accidents and poisoning

Ill defined conditions (old age)

Unspecified

All causes

Appendix A

Table 1. Aggregation of causes.

Group	Condition	Condition(s) in Table 3
1	Infectious and parasitic diseases	1-44
2	Neoplasm (Cancer)	45-61
3	Endocrine, nutritional and metabolic disease	62-66
4	Blood disorders and blood forming organs	67-68
5	Mental disorders	69-71
6	Disorders of the nervous system	72-79
7	Diseases of the circulatory system	80-88
8	Diseases of the respiratory system	89-96
9	Diseases of the digestive system	97-104
10	Genitourinary disorders	105-111
11	Complications of pregnancy, child birth and the puerperium	112-118
12	Skin and sub-tissue disorders	119-120
13	Diseases of the musculoskeletal system	121-125
14	Congenital anomalies	126-130
15	Prenatal disorders	131-135
16	Accidents and poisoning	138-150
17	Senility without mention of psychosis (old age)	136
18	Unspecified	137+151

Appendix B

Table 2: Mortality causes.

1. Cholera	25. Measles
2. Typhoid fever	26. Yellow fever
3. Paratyphoid fever & other salmonela infections	27. Viral encephalitis
4. Bacillary dysentery and amoebiasis	28. Infectious hepetitis
5. Enteritis and other diarrhoeal diseases	29. Other viral diseases
6. Tuberculosis of respiratory system	30. Typhus and other rickettsioses
7. Tuberculosis of meninges and central nervus system	31. Malaria
8. Tuberculosis of intestines peritoneum and mesenteric glands	32. Trypanosomiasis
9. Tuberculosis of bones and joints	33. Relapsing fever
10. Other tuberculosis including late effects	34. Congenital syphilis
11. Plague	35. Early syphilis, symptomatic
12. Anthrax	36. Syphilis of cenral nervous system
13. Brucellosis	37. Other syphilis
14. Leprosy	38. Gonococcal infections
15. Diphtheria	39. Schistosomiasis
16. Whooping cough	40. Hydatidosis
17. Streptococcal sore throat and acarlet fever	41. Filarial infection
18. Erysipelas	42. Ancylostomiasis
19. Meningococcal infection	43. Other helminthiasis
20. Tetanus	44. All other infective and parasitic diseases
21. Other bacterial diseases	45. Malignant neoplasm of buccal cavity & pharynx
22. Acute poliomyelitis	46. Malignant neoplasm of oesophagus
23. Late effect of acute poliomyelitis	47. Malignant neoplasm of stomach
24. Smallpox	48. Malignant neoplasm of intestine, except rectum
	49. Malignant neoplasm of rectum & rectosigmoid

junction	102. Cirrhosis of liver
50. Malignant neoplasm of larynx	103. Cholelithiasis and cholecystitis
51. Malignant neoplasm of trachea, bronchus & lung	104. Other diseases of digestive system
52. Malignant neoplasm of bone	105. Acute nephritis
53. Malignant neoplasm of skin	106. Other nephritis and nephrosis
54. Malignant neoplasm of breast	107. Infections of kidney
55. Malignant neoplasm of cervix uteri	108. Calculus of urinary system
56. Other malignant neoplasm of uterus	108. Hyperplasia of prostate
57. Malignant neoplasm of prostate	109. Diseases of breast
58. Malignant neoplasm of other and unspecified sites	111. Other diseases of genito-urinary system
59. Leukaemia	112. Toxaemias of pregnancy and the puerperium
60. Other neoplasm of lymphatic & haemotopoietic tissue	113. Haemorrohage of pregnancy and childbirth
61. Benign neoplasm & neoplasms of unspecified nature	114. Abortion induced for legal indications
62. Non-toxic goitre	115. Other and unspecified abortion
63. Thyrotoxicosis with or without goitre	116. Sepsis of childbirth and the puerperium
64. Diabetes mellitus	117. Other complication of pregnancy, childbirth and the puerperium
65. Avitaminoses and other nutritional deficiency	118. Delivery without mention of complication
66. Other endocrine and metabolic diseases	119. Infections of skin and subcutaneous tissue
67. Anaemias	120. Other diseases of skin and subcutaneous tissue
68. Other diseases of blood and blood-forming organs	121. Arthritis & spondylitis
69. Psychoses	122. Non-articular rheumatism and rheumatism unspecified
70. Neuroses, personality disorders & other non-psychotic mental disorders	123. Osteomyelitis and periostitis
71. Mental retardation	124. Ankylosis and acquired musculoskeletal deformities
72. Meningitis	125. Other diseases of musculoskeletal system and connective tissue
73. Multiple sclerosis	126. Spina bifidia
74. Epilepsy	127. Congenital anomalies of heart
75. Inflammatory diseases of eye	128. Other congenital anomalies of circulatory system
76. Cataract	129. Cleft palate and cleft lip
77. Glaucoma	130. All other congenital anomalies
78. Otitis media and mastoiditis	131. Birth injury and difficult labour
79. Other diseases of nervous system and sense organs	132. Conditions of placenta and cord
80. Active rheumatic fever	133. Haemolytic disease of newborn
81. Chronic rheumatic heart disease	134. Anoxic and hypoxic conditions not elsewhere classified
82. Hypertensive disease	135. Other causes of prinatal morbidity and mortality
83. Ischaemic heart disease	136. Senility without mention of psychosis
84. Other forms of heart disease	137. Symptoms and other ill-defined conditions
85. Cerebrovascular disease	138. Motor vehicle accidents
86. Diseases of arteries arterioles & capillaries	139. Other transport accidents
87. Venous thrombosis and embolism	140. Accidental poisoning
88. Other diseases of circulatory system	141. Accidental falls
89. Acute respiratory infection	142. Accidents caused by fires
90. Influenza	143. Accidental drowning and submersion
91. Viral pneumonia	144. Accidents caused by firearm missiles
92. Other pneumonia	145. Accidents mainly of industrial type
93. Bronchitis, emphysema and asthma	146. All other accidents
94. Hypertrophy of tonsils and adenoids	147. Suicide and self inflicted injury
95. Empyema and abscess of lung	148. Homicide and injury purposely inflicted by intervention
96. Other diseases of respiratory system	149. Injury undetermined whether accidentally or purposely inflicted
97. Diseases of teeth and supporting structures	150. Injury resulting from operations of war
98. Peptic ulcer	151. Unspecified causes
99. Gastritis and duodenitis	
100. Appendicitis	
101. Intestinal obstruction and hernia	

References

Aalabaf-Sabaghi, M. (2002) Mortality Risks in Iran, E.C.O. College of Insurance Discussion Paper.

Benjamin, D.K. and W.R. Dogan (1997) Individuals Estimates of the Risks of Death: Part I- A Reassessment of the Previous Evidence, *Journal of Risk and Uncertainty,* 15, 115-133.

Benjamin, D.K., W.R. Dogan and D. Buschena (2001) Individuals' Estimates of the Risk of Death: Part II- New Evidence, *Journal of Risk and Uncertainty,* 22:1; 35-57.

Canals-Cerda, J. and S. Gurmu (2007) "Semiparametric competing risks analysis", *Econometrics Journal*, 10, 193-215

Corro, Dan, (2001) "A Note on the Inverse Relationship Between Hazard and Life Expectancy", *Casualty Actuary Society, Forum*, fall, 475-492.

Cox, D. R. (1972) "Regression Models and Life Tables", *Journal of the Royal Statistical Society*, Series B, 2, 34,187-200.

Elbers, C. and G. Ridder (1982) True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model, *Review of Economic Studies*, 49, 403-411.

Grossman, M. and L. Benham (1974) "Health, Hours and Wages", in M. Perlman ed., The Economics of Health and Medical Care, Macmillan, London.

Heckman, J. and B. Singer (1984a) "The Identifiability of the Proportional Hazard Model", *Review of Economic Studies*, 51, 231-241.

-------------------------------- (1984b) A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, *Econometrica*, 52, 271-320.

Le, C.T. (1997) Applied Survival Analysis, Wiley.

London, D. (1997) "Survival Models and Their Estimation", Actex Publications.

Newman, J.L. and C.E. McCulloch (1984) A Hazard Rate Approach to the Timing of Birth, *Econometrica*, 52, 4,939-962.

Sickles, R.C. and Taubman, P. (1997) Mortality and Morbidity Among Adults and the Elderly, in Handbook of Population and Family Economics, editors, M.R. Rosenzweig and O. Stark, NorthHolland.

Sunstein, C.R. (1997) Bad Deaths, *Journal of Risk and Uncertainty,* 14, 259-282.

Viscusi, W.K., J.K. Hakes and A. Carlin (1997) Measures of Mortality Risks, *Journal of Risk and Uncertainty,* 14:213-233.

Zanjani, H. and T. Noorallahi (1999) The Iranian Life Table for 1996, The Higher Institute for Social Security Research, Tehran, Iran.