PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1189 — 2007
Zastosowania metod ilościowych

Zofia Mielecka-Kubień
Akademia Ekonomiczna w Katowicach

ESTIMATION OF LIFE TABLES PARAMETERS FOR SMOKERS BASED ON RELATIVE RISK

1. General remarks

One of the most important negative effects of smoking is premature mortality of smokers (in comparison to mortality of non-smokers or to the one in general population). The problem has been often considered[1] and many measures (e.g.: relative risk, odds ratio, relative hazard, excess risk, attributable risk) showing the difference between smokers and non-smokers' mortality can be applied for the purpose. One of the possibilities is to estimate life tables parameters for smokers and to compare them with the parameters estimated for no-smoking (or general) population. In such case the life tables for the no-smoking (or general) population could provide a "baseline" for the comparison with the life tables parameters for smokers. The problem is that the number of deaths related to smoking (needed to obtain the basic input data for the estimation of the life tables parameters – that is the probability of deaths or the mortality rates for smokers) is very difficult to evaluate. In general "*...The matter* [of stating the cause of death] *is compounded by classification procedures which require assignment of a single underlying 'cause' of death, implying presence or absence of the condition without allowance for uncertainty...*" [Gittelsohn 1982]; in the case of smoking – as the relation between smoking and mortality is indirect – the task is even more difficult, as it is difficult to evaluate, even roughly, the range of smoking contribution to the death of a smoker. The author's idea is to base the estimation of the probability of deaths of smokers (or similarly mortality rates for smokers) on relative risk, comparing mortality[2] of a population exposed to the risk factor (smoking) with the one unexposed to the factor. In that way the troublesome and controversial estimation of the range of the influence of smoking[3] on the deaths of smokers can be avoided.

---

[1] See [Godley 1975; *The Economic...*1986; Collishaw et al. 1988; Breslow 1996], and many others.

[2] Relative risk is estimated for mortality as well as for morbidity related to exposure to a risk factor.

[3] See for instance [*The Economic...*1986].

Let

D – denotes a cause of death attributable to smoking,

$\overline{D}$ – the other causes of death,

E, $\overline{E}$ – exposure and non-exposure to the risk factor (smoking).

Relative risk γ* (as an approximation of exposure odds ratio) can be written as [Breslow 1996, p. 15].

$$\gamma^* = \frac{P(E/D)\cdot P(\overline{E}/\overline{D})}{P(E/\overline{D})\cdot P(\overline{E}/D)} = \frac{P(D/E)\cdot P(\overline{D}/\overline{E})}{P(D/\overline{E})\cdot P(\overline{D}/E)} \approx \frac{P(D/E)}{P(D/\overline{E})}. \tag{1}$$

Relative risk γ^* must be a nonnegative number; $\gamma^* = 1$ denotes that $P(D/E) = P(D/\overline{E})$ which means that D and E are independent. If $\gamma^* > 1$, then there is greater risk or probability of D when exposed (E) then unexposed ($\overline{E}$); if $\gamma^* < 1$ the situation is reverse. Since $P(D/E)$ can be no larger than 1, γ^* must be less or equal to $1/P(D/\overline{E})$.

To present Mantel-Haenszel relative risk estimator[4] γ the following notation is introduced:

Table 1. Notations for Mantel-Haenszel relative risk estimator

Observed frequencies	Cause of death			
Exposure	Notation	*D*	*not D*	Total
	E	a_i	b_i	a_i+b_i
	not E	c_i	d_i	c_i+d_i
	Total	a_i+c_i	b_i+d_i	n_i

Source: [Jewell 2004, p. 124].

Mantel-Haenszel estimator can be written as

$$\gamma = \frac{\sum_i (a_i d_i / n_i)}{\sum_i (b_i c_i / n_i)} \tag{2}$$

with the variance estimator, where $\beta = \log\gamma$

$$\mathrm{var}(\beta) = \frac{1}{\left(\sum_i (a_i d_i / n_i)\right)^2} \sum_i n_i^{-2}(a_i d_i + \gamma b_i c_i)\cdot[a_i + d_i + \gamma(b_i + c_i)]. \tag{3}$$

[4] See [Jewell 2004, pp.124-130]. The estimator is constructed for population stratified (i – denotes a stratum) for instance according to gender and age, or any other confounding factor.

In practice[5] the approximate value of relative risk is often estimated (see formulae 1) as

$$\gamma \approx \frac{\sum_i [a_i/(a_i+b_i)]}{\sum_i [c_i/(c_i+d_i)]}. \tag{4}$$

Relative risk γ is easy to interpret – if for instance the relative risk of lung cancer morbidity attributable to smoking is 10.73 [www.BiologyPages/E/Epidemiology], a smoker's lifetime risk of lung cancer increases over 10 times.

2. The estimation of the probability of deaths of smokers

Because of the usually observed shortcomings of statistical data, it has been assumed that population consists exclusively of smokers and non-smokers.

Let

p, p_s, p_n – denote respectively probability of deaths, probability of deaths of smokers, and of non-smokers in the considered population in the given year,

l_s, l_n – number of smokers and non-smokers,

L – population number, $l_s + l_n = L$,

$w_s = \frac{l_s}{L}$ – share of smokers in the population,

$w_n = \frac{l_n}{L}$ – share of non-smokers in the population, $w_s + w_n = 1$,

$z_s = p_s \cdot l_s$ – number of deaths of smokers,

$z_e = p_n \cdot l_s$ – expected number of deaths of smokers,

$\gamma = \frac{p_s}{p_n} = \frac{p_s \cdot l_s}{p_n \cdot l_s} = \frac{z_s}{z_e}$ – relative risk for smokers equal to observed/expected number of deaths of smokers.

The probability of death in a population can be written as the weighted average of the probability of death of a smoker and that of a non-smoker, where weights are shares (w_s, w_n) of both subpopulations:

$$p = w_s \cdot p_s + w_n \cdot p_n, \tag{5}$$

as $w_s + w_n = 1$ $\qquad p = w_s \cdot p_s + (1 - w_s) \cdot p_n,$ (6)

$p_n = \frac{p_s}{\gamma}$ so $\qquad p = w_s \cdot p_s + (1 - w_s) \cdot \frac{p_s}{\gamma}.$ (7)

[5] See e.g. [Jewell 2004, p. 32].

From (7) one can obtain the estimate of the probability of death of a smoker

$$p_s = \frac{p \cdot \gamma}{w_s(\gamma - 1) + 1}, \tag{8}$$

and the estimate of the probability of death of a non-smoker

$$p_n = \frac{p}{w_s(\gamma - 1) + 1}. \tag{9}$$

3. Empirical results

Smoking prevalence in Poland in the last several years is, generally speaking, decreasing (fig. 1 and 2); in 2003 it was equal to 36.9% for men and 21.1% for women (population aged 15 years and older). The share of smokers in Poland is much higher in the population of men than that of women, but the difference between the two populations tends to diminish (fig. 2).

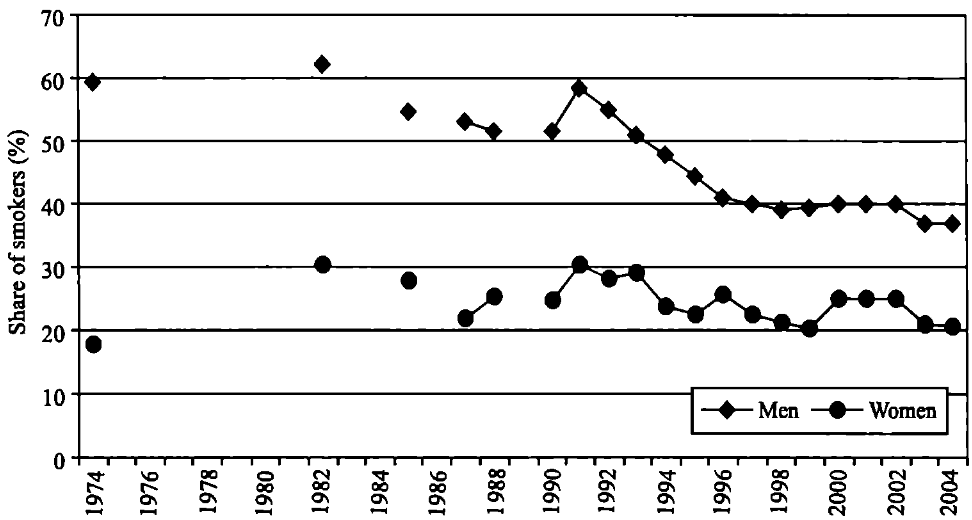

Fig. 1. Share of current smokers (%) in Poland in 1974, 1982, 1985, 1987-1988, 1990-2004 (percent of male and female population aged 15 years and older)

Source: author's own on the basis of OBOP (1974-1994) and Global Market Information Database [www.gmid.euromonitor.com] (1995-2004).

As the estimation of the smoking-attributable mortality requires data on relative risk as well as on smoking prevalence, and the both vary considerably according to age, it is advisable to control both phenomena according to age through examining smoking-attributable mortality by 5-year age classes. The probability of death of a smoker can be then estimated as

$$p_{si} = \frac{p_i \cdot \gamma_i}{w_{si}(\gamma_i - 1) + 1}, \tag{10}$$

where i – denotes gender/age class.

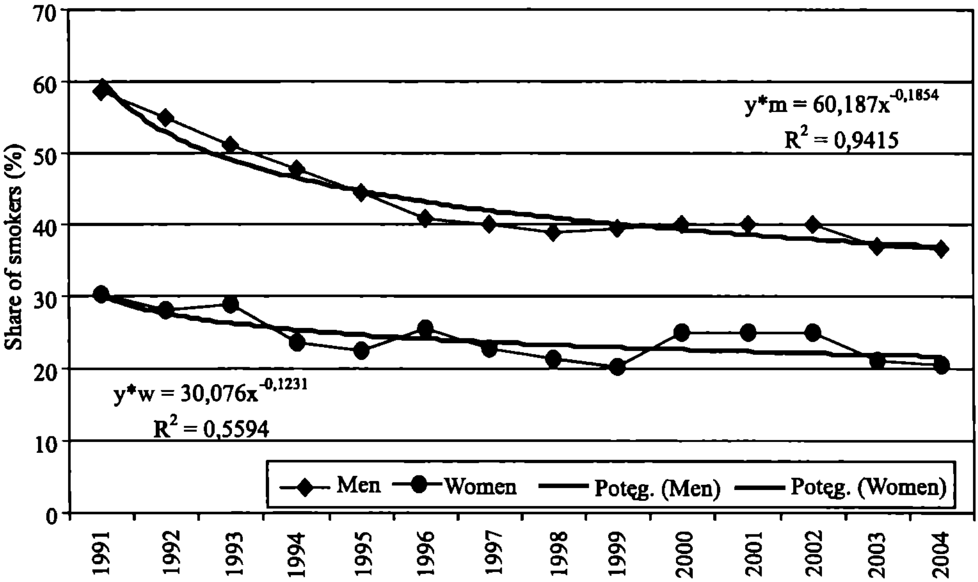

Fig. 2. Share of current smokers in Poland in 1991-2004[6]
(percent of male and female populations aged 15 years and older)
Source: author's own on the basis of OBOP (1991-1994) and Global Market Information Database [www.gmid.euromonitor.com] (1995-2004).

Table 2. Number of daily smokers according to gender and age in Poland, 1996 (thousands)

Age	Men	Women
15-24	677,9	298,3
25-44	2935,2	1801,6
45-64	1718,5	788,0
65-74	324,7	87,9
75+	58,5	10,9
Total	5724,8	2986,7

Source: [*Stan*... 1999, pp. 112-117].

Relative risk for smokers was never estima-ted for the population of Poland, therefore the results of the foreign prospective epidemiolo-gical study have to be applied. The F.H. God-ley's set of risk estimators was used[7], and to avoid some random fluctuations, the theoretical values of fitted curves were applied (fig. 3).

To estimate the smoking prevalence according to gender/age classes the results of health status survey in Poland (1996)[8] presented in table 2 were applied.

[6] Presented models have only illustrative purpose, therefore they were not further analyzed.
[7] [Godley 1975], quoted after [Collishaw et al. 1988]. It is one of few available studies where relative risk is estimated in 5-year classes.
[8] [*Stan*...1999]. The survey is repeated every few years.

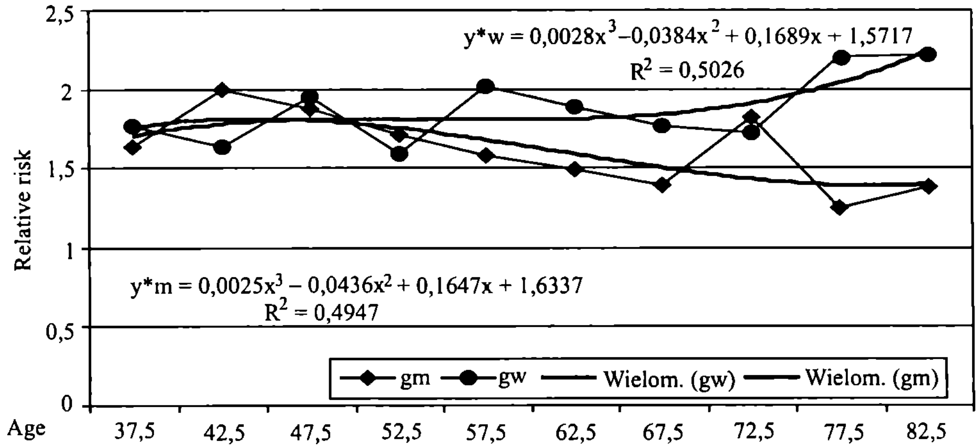

Fig. 3. Empirical and theoretical values of relative risk for men (gm) and women (gw)

Source: author's own on the basis of [Godley 1975].

To split the number of smokers into 5-year classes it has been assumed that the distribution of the number of smokers according to age can be described with a log-normal curve. The supposition is supported by the approximately linear relation between the probits and the logarithms of age in male and female subpopulations (fig. 4 and 5). Figure 6 presents the estimated numbers of daily smokers in 5-year classes.

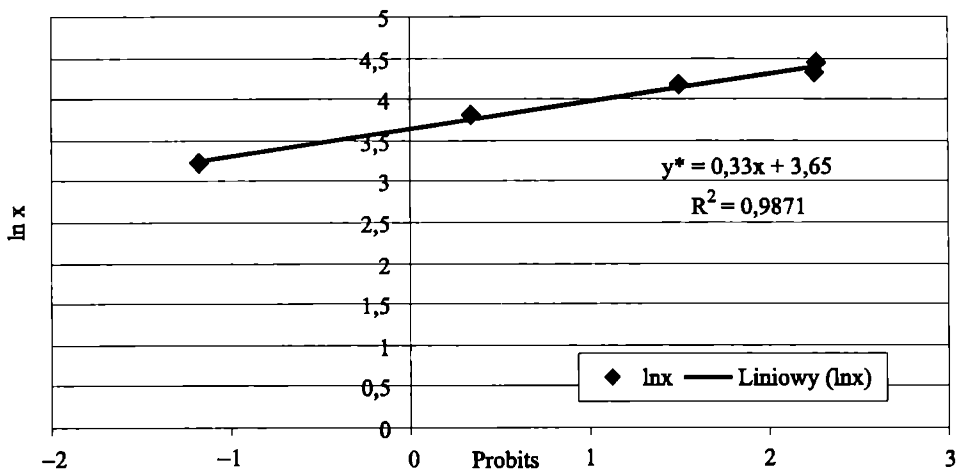

Fig. 4. Relation between probits and logarithms of age, men

Source: author's own on the basis of [*Stan…* 1999].

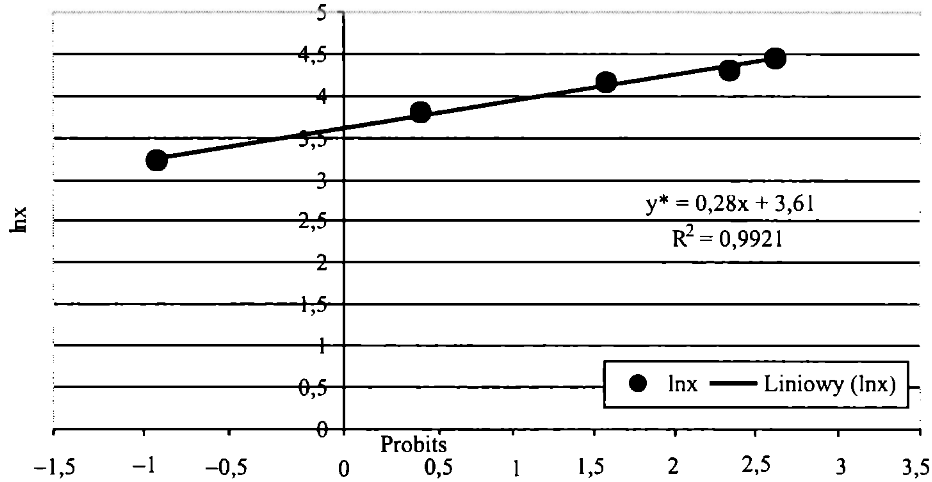

Fig. 5. Relation between probits and logarithms of age, women

Source: author's own on the basis of [*Stan*... 1999].

On that basis, for the Polish population 2003, the probability of death of smokers in gender/age class *i* (according to formulae /10/), was estimated and compared with the probability of death of a non-smoker (formulae /9/) and in the general population (fig. 6 and 7). Next, there were estimated the life tables parameters for smokers, non-smokers and the whole population; the results are presented in tables 3 and 4.

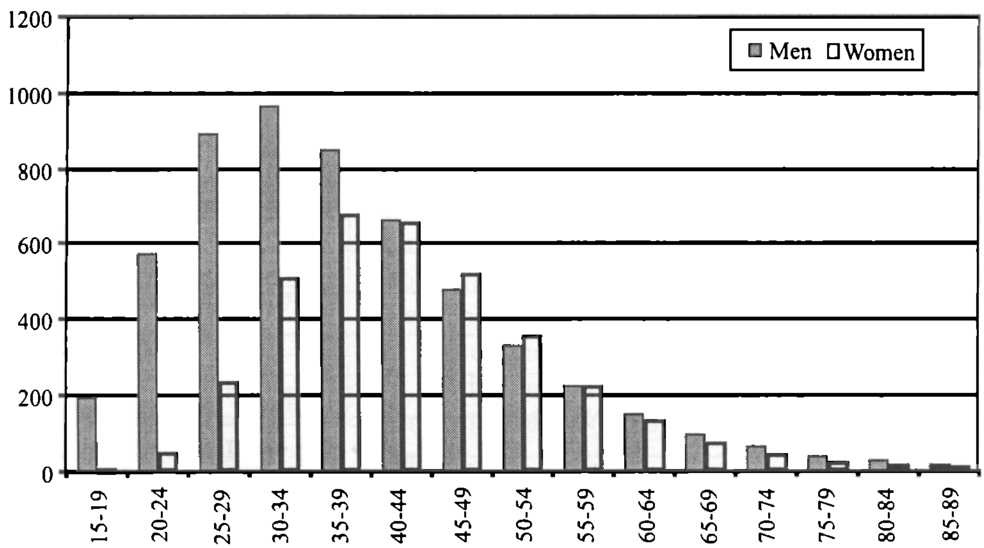

Fig. 6. Estimated numbers of daily smokers in 5-year classes

Source: author's own on the basis of [*Stan*... 1999].

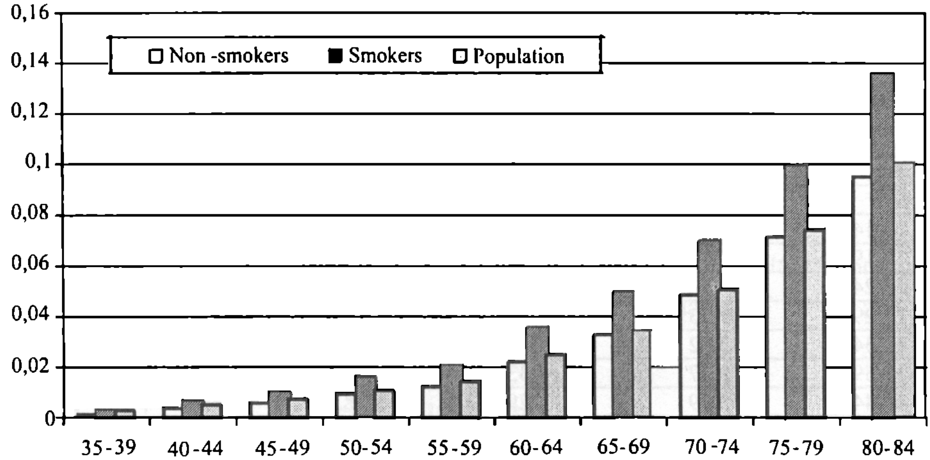

Fig. 7. Estimated probability of death of smokers, non-smokers and population, men

Source: author's own on the basis of [*Stan*... 1999; Godley 1975] and *Demographic Yearbook of Poland 2003*.

It can be observed that there are significant differences among the probabilities of deaths in the considered subpopulations. The probability of death of smokers is much higher than that of non-smokers or in general population; the difference is growing with age, which is caused by the time-lag between the 'exposure to smoking' (as one of the causes of death) and the moment of death. As the probabilities of death in female general and no-smoking subpopulations are lower than those of men, the mentioned above differences are bigger for females.

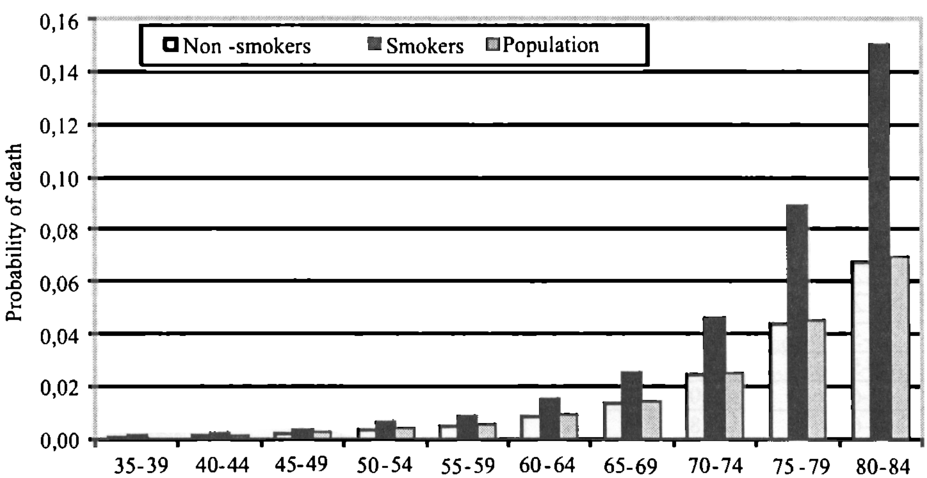

Fig. 8. Estimated probability of death of smokers, non-smokers and population, women

Source: author's own on the basis of [*Stan*... 1999; Godley 1975] and *Demographic Yearbook of Poland 2003*.

Table 3. Estimated life expectation (ex) for non-smokers, smokers, population, and the differences between smokers and non-smokers (s-n), smokers and population (s-p), men

Age	Ex			Difference	
	Non-smokers	Smokers	Population	s-n	s-p
0	71,39	66,80	70,13	–4,59	–3,33
1-4	70,98	66,36	69,71	–4,62	–3,35
5-9	67,08	62,45	65,81	–4,63	–3,36
10-14	62,15	57,52	60,88	–4,63	–3,36
15-19	57,23	52,59	55,95	–4,64	–3,36
20-24	52,44	47,78	51,16	–4,66	–3,38
25-29	47,73	43,04	46,44	–4,69	–3,40
30-34	43,02	38,31	41,73	–4,71	–3,42
35-39	38,37	33,62	37,07	–4,75	–3,45
40-44	33,69	29,10	32,53	–4,59	–3,43
45-49	29,22	24,92	28,24	–4,30	–3,32
50-55	24,99	21,10	24,17	–3,89	–3,07
55-59	21,04	17,66	20,38	–3,38	–2,72
60-64	17,22	14,31	16,71	–2,91	–2,40
65-69	13,96	11,61	13,60	–2,35	–1,99
70-74	11,00	9,18	10,73	–1,82	–1,55
75-79	8,34	7,00	8,12	–1,34	–1,12
80+	5,85	4,97	5,67	–0,88	–0,70

Source: author's own on the basis of [*Stan*... 1999; Godley 1975] and *Demographic Yearbook of Poland 2003, 2004.*

Table 4. Estimated life expectation (ex) for non-smokers, smokers, population, and the differences between smokers and non-smokers (s-n), smokers and population (s-p), women

Age	Ex			Difference	
	Non-smokers	Smokers	Population	s-n	s-p
0	79,17	74,26	78,46	–4,91	–4,20
1-4	78,71	73,78	78,00	–4,93	–4,22
5-9	74,80	69,85	74,09	–4,95	–4,24
10-14	69,86	64,91	69,14	–4,95	–4,23
15-19	64,91	59,96	64,20	–4,95	–4,24
20-24	60,00	55,04	59,29	–4,96	–4,25
25-29	55,09	50,12	54,37	–4,97	–4,25
30-34	50,17	45,19	49,45	–4,98	–4,26
35-39	45,28	40,30	44,57	–4,98	–4,27
40-44	40,42	35,50	39,75	–4,92	–4,25
45-49	35,64	30,85	35,05	–4,79	–4,20
50-55	30,99	26,38	30,48	–4,61	–4,10
55-59	26,49	22,14	26,07	–4,35	–3,93
60-64	22,07	18,02	21,74	–4,05	–3,72
65-69	17,89	14,21	17,65	–3,68	–3,44
70-74	13,99	10,78	13,80	–3,21	–3,02
75-79	10,46	7,92	10,31	–2,54	–2,39
80+	7,40	5,98	7,27	–1,42	–1,29

Source: author's own on the basis of [*Stan*... 1999; Godley 1975] and *Demographic Yearbook of Poland 2003, 2004.*

According to presented estimation results (tables 3 and 4) life expectancy loss

$$e_x^l = e_x^n - e_x^s \quad (11)$$

or

$$e_x^l = e_x - e_x^s \quad (12)$$

where e_x^l, e_x^n, e_x^s, e_x denote respectively life expectancy loss, life expectancy for smokers, non-smokers and in general population, is nearly equal to 5 years and is higher for females than for males.

4. Conclusions

Smoking definitely shortens the life expectation of the Polish smoking subpopulation. The estimated life expectancy loss is slightly bigger for women. Applying the relative risk estimators enables to estimate the probability of death of smokers and their life tables parameters without very difficult and controversial evaluation of smoking contribution to the causes of death of smokers. The accuracy of the presented estimation would be higher, if the relative risk and the smoking prevalence were estimated for the Polish population in at least 5-year long gender/age classes.

References

Breslow N.E., *Statistics in Epidemiology: The Case-Control Study*, „Journal of the American Statistical Association" 1996, pp. 14-28.

Collishaw N.E., Tostowaryk W., Wigle D.T., *Mortality Attributable to Tobacco Use in Canada*, „Canadian Journal of Public Health" 1988, pp.166-169.

Demographic Yearbook of Poland 2003, 2004, GUS, Warsaw.

Gittelsohn A.M., *On the Distribution of Underlying Causes of Death*, American „Journal of Public Health"1982, pp.133-140.

Godley F.H., *Cigarette Smoking, Social Factors, and Mortality: New Estimates from Representative National Samples*, Ann Arbor, Michigan (university microfilm) 1975.

Jewell N.P., *Statistics for Epidemiology*, Chapman&Hall/CRC, Boca Raton, London, New York, Washington D.C.2004.

Stan zdrowia ludności Polski w przekroju terytorialnym w 1996 r., GUS, Warszawa 1999.

The Economic Consequences of Smoking in Northern Ireland, Ulster Cancer Foundation 1986.

www.BiologyPages/E/Epidemiology.

www.gmid.euromonitor.com.

OSZACOWANIE PARAMETRÓW TABLIC TRWANIA ŻYCIA DLA PALACZY Z WYKORZYSTANIEM WSPÓŁCZYNNIKÓW RYZYKA WZGLĘDNEGO

Streszczenie

Do najważniejszych negatywnych skutków palenia tytoniu zaliczyć można przedwczesną umieralność palaczy. W artykule zaprezentowano propozycję autorki szacowania parametrów tablic trwania życia dla palaczy z wykorzystaniem współczynników ryzyka względnego (*relative risk*), obliczanego jako stosunek prawdopodobieństwa zgonu palacza do prawdopodobieństwa zgonu niepalacza. To podejście pozwala uniknąć przy oszacowaniu parametrów tablic trwania życia, a także prawdopodobieństw zgonu palaczy i niepalaczy w badanej populacji, trudnej i kontrowersyjnej oceny znaczenia palenia tytoniu dla określonej przyczyny zgonu palacza.