Modeling melanoma incidence trends

Örjan Hallberg, Hallberg Independent Research, Sweden

Introduction

A statistical model was earlier developed with the aim of predicting future melanoma rates in different countries [1-3]. The last step in this modeling work was published April 2008 by European Journal of Cancer Prevention [3]. We will here give a more detailed description of the model and its practical use, so that it might be of help for others in analyzing melanoma trends in more countries.

Traditionally, authorities and epidemiologists have focused on the skin damage generating mechanism from UV exposure when the increasing melanoma rates are being addressed. The hypothesis is then that the body's natural capability to repair or to kill damaged skin cells is not large enough to cope with the higher number of skin damages attained due to increased exposure to UV radiation.

There is, however, a possibility that the increasing melanoma rates are due to reduced repair efficiency rather than due to overwhelming damage rates. The model developed makes it possible to vary cancer risk over time, repair rate over time and the skin damage pattern over life for a whole population. It may also be possible to request an Excel application for own evaluation, use and further development.

Model principles

One birth cohort

The key point of the model is that we are looking at damages attained during each one of the life years of a birth cohort separately. The damages collected e.g. during the 15th year of a person's life are followed for the remaining life time quite separately from the damages collected during the next year etc. Table I shows the principle lay-out of the calculation matrix for one birth cohort.

		Year								
Year	*Dam.*	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	.	.
Y_1	N_1	N_1F_1	N_1F_2	N_1F_3	N_1F_4	N_1F_5	N_1F_6	N_1F_7	.	.
Y_2	N_2		N_2F_1	N_2F_2	N_2F_3	N_2F_4	N_2F_5	N_2F_6	.	.
Y_3	N_3			N_3F_1	N_3F_2	N_3F_3	N_3F_4	N_3F_5	.	.
Y_4	N_4				N_4F_1	N_4F_2	N_4F_3	N_4F_4	.	.
Y_5	N_5					N_5F_1	N_5F_2	N_5F_3	.	.
Y_6	N_6						N_6F_1	N_6F_2	.	.
Y_7	N_7							N_7F_1	.	.

		? Y_1	? Y_2	? Y_3	? Y_4	? Y_5	? Y_6	? Y_7		

Table I. Skin damages attained during year *i* (N_i) contribute to the melanoma risk by N_iF_j . The total risk per year is given by the vertical sum ? $\mathbf{Y_i}$

The risk function F_j is defined as an increasing melanoma risk function M_i multiplied by a decreasing repair function R_j representing remaining level of active, unrepaired skin damages left. The repair function is assumed to initially follow a natural, efficient path, capable of repairing most of skin damages in due time. However, the model also opens up for the possibility that this repair efficiency suddenly, or gradually, becomes reduced leaving more damages left old enough for the melanoma risk function to reach dangerous high levels. The detailed mathematics for these calculations can be reviewed directly in the Excel application.

Age standardized rates

In order to calculate age-standardized rates we need calculated or reported data from all age groups. Reported data are available from cancer registries in many countries, both age-specific and age-standardized data. For calculation of age-standardized data based on calculated age-specific data we need to create a series of matrixes like in Table I but representing birth cohorts from 1860 up to 1980. This makes it possible to calculate the effect of a sudden or gradually reduced repair efficiency as it affects different age-groups from the same point in calendar time.

These calculated age-specific rates are then combined to form the calculated world age standardized rate year by year. Since the basic models used only contain a few parameters it is possible to fit the calculated age-standardized rates to reported rates by varying only 2 parameters.

Verification test

Once the parameters have been optimized to fit reported age-standardized data it is easy to plot the calculated age-specific rates over calendar time to compare with reported data. If reported and calculated data fits well this is in support of the basic model. If not, the basic principles of the model should be reviewed once more.

Results

Results from using this method and the Excel application are presented in ref. [3]. In summary, the model was tested using data from three countries for which the two parameters used for data fitting became quite similar. In order to test the commonly assumed main cause to increasing melanoma rates we introduced an increasing volume of skin damaging UV exposure over time, leaving the repair efficiency unchanged at the natural level. After best fit to age-standardized data it appeared that the fit to age-specific data was very poor. In conclusion we found that the reduced repair hypothesis had a much better capability of explaining past and current trends than the ´sun shine´ hypothesis had.

Description of the Excel application

The main calculation sheet, "Calc" contains all matrixes used for the detailed calculation of age-specific incidence. On top of that sheet it is possible to view calculated and reported incidence over age for any wanted birth cohort. As an example Figure 1 shows this for the population born in 1925 in Sweden. It also illustrates the trend-breaks in repair efficiency after 1955 for skin damages collected in different years by this birth cohort.

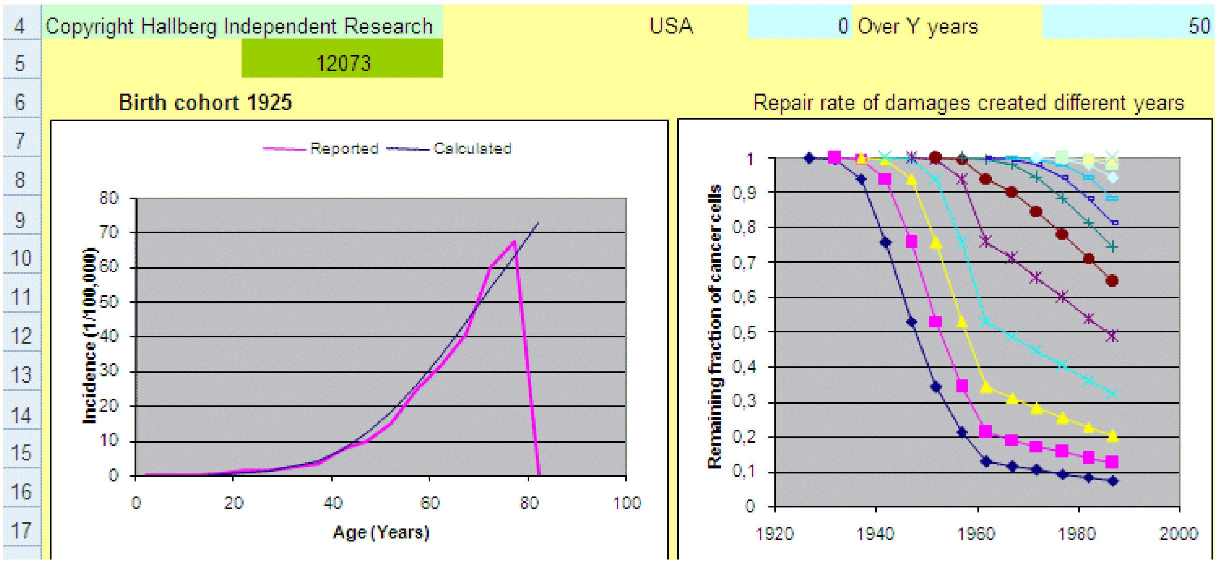


Figure 1. Calculated and reported incidence in Sweden for people born in 1925. The trend break in repair efficiency from 1955 is shown to the right.

Statistical distributions

The statistics used for cancer risk over time and repair rate over time were based on log-normal distributions using only 2 parameters for each distribution. One parameter was time to 0.1% accumulated, used instead of time to 50% (median time). The other was dispersion, here expressed in number of time decades. The by far most sensitive parameter for this study was the dispersion, while the functions could be nailed to pass through 0.1% at 100 years for cancer risk and 5 years for 0,1% repair probability. By varying the dispersions only, all forms of curves could be found including the one giving best fit to reported data.

Figure 2 shows the natural and the disturbed repair probability over time. The natural repair was characterized by a dispersion of 0.2 decades that was used for all countries while the dispersion of the disturbed repair probability was varied to fit reported data, as shown later on.

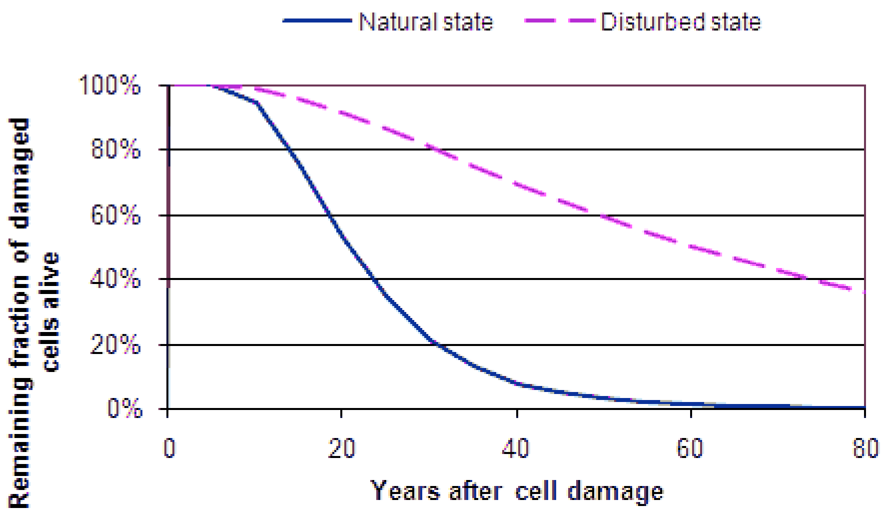


Figure 2. The graph shows the remaining fraction of damaged cells still unrepaired and alive for the natural state and in the disturbed state according to Swedish data.

If no damage repair took place after initiation, the accumulated melanoma risk after one year of collected skin damages would follow the path shown in Figure 3. As stated before, this line was set to pass 0.1% at 100 years so the shape of the curve was entirely controlled by one parameter only, the dispersion.

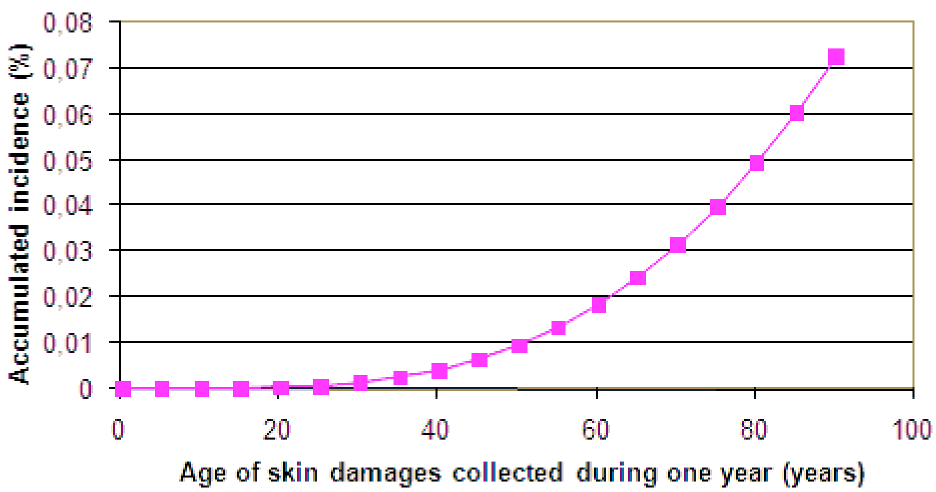


Figure 3. The risk that skin damages collected during one year might develop into melanoma was described by this function for the Swedish population for the case no repairs took place.

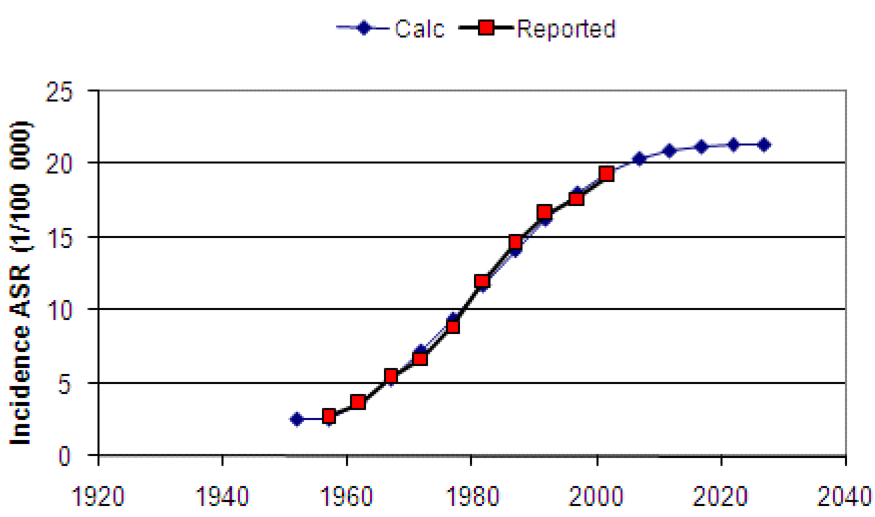
The combined effect of the functions in Figures 2 and 3 were shown in ref. 3 but can also be found directly from the Excel application MelIncTool.

Parameter optimization

Before any curve fitting can take place you need to supply reported data to the application. In the current state it has been preloaded with data from Sweden, Norway and from the USA. On top of the “Calc”-sheet you can define which one of the three sets of data you want to analyze. These data sets can easily be replaced by your own data if you would like to.

Reported age-standardized data over calendar time are used to determine parameters giving the best fit to reported data. Figure 4 shows that only two parameters are varied while the rest are fixed.

0,077973
0,060369
0,054845
0,058238
0,065171
0,064515
0,062945
0,059547
0,049842
0,038964
0,027212
0,015954
0,006925



Sweden	Norway	USA				Dispersioı	t0.1% Yr
0,462057	0,492342	0,506426	100		F(t)	0,462057	100
0,349423	0,367046	0,491282	5		S(t) dist.	0,349423	5
0,2	0,2	0,2	5		S(t) nat.	0,2	5

Figure 4. The graph shows how the two parameters in green field have been varied to make calculated age-standardized incidence to fit reported data. For Sweden we used ASR 1970 while the world standard was used for Norway and the USA.

The standard tool for problem solving provided by Excel was used for the application.

Evaluation of the model correctness

Even if the calculated age-standardized rates fit reported data perfectly as in Figure 4, there is no guarantee that also corresponding age-specific data will fit reported data. Different combinations of incidence rates from young and old people might give the same total sum when the age-standardized rates are calculated. Therefore it is important to check the model by comparing also calculated and reported age-specific rates. The MelIncTool automatically gives these two graphs so you can judge for yourself. See Figures 5 and 6.

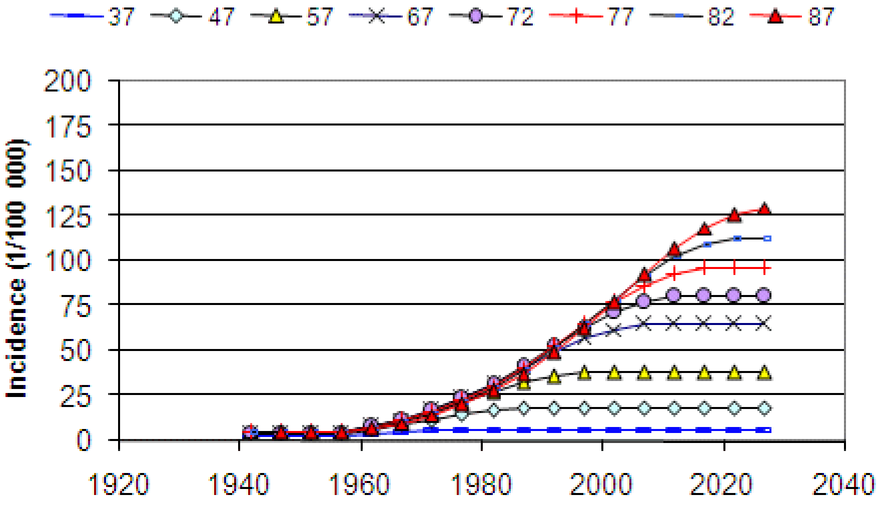


Figure 5. Calculated age-specific incidence in Sweden levels off at increasing levels for increasing ages. Each symbol represents a specific age defined on top of the graph.

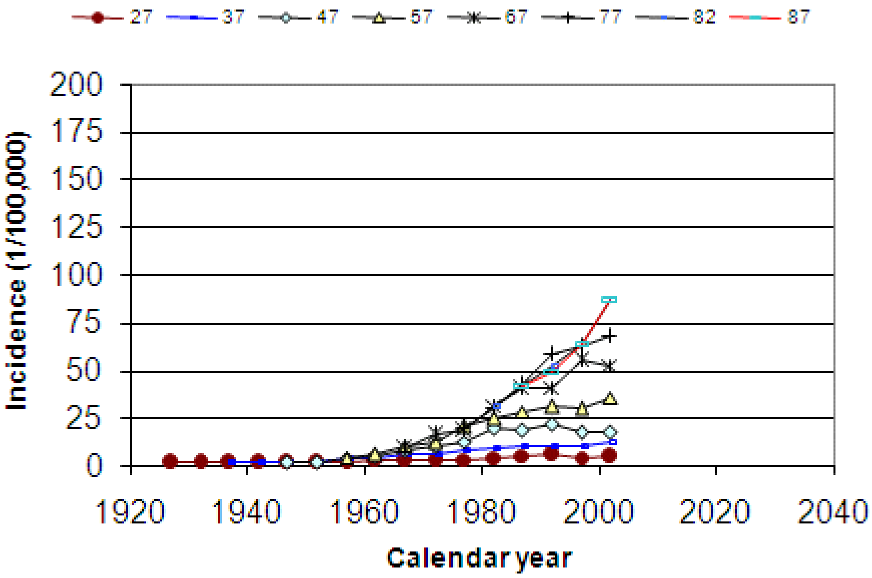


Figure 6. This graph gives officially reported melanoma incidence numbers for men in Sweden.

It is obvious that the two graphs above actually fit very well, thus supporting the relevance of the model used.

It is also possible to compare calculated and reported incidence trends for persons born in the first half of the 20th century. These people first had very low incidence rates corresponding to a good, natural repair capability. But right after 1955 all these birth cohorts suddenly started to get melanoma, independent of year of birth. See Figure 7. The MelIncTool gives these data automatically when a country has been selected.

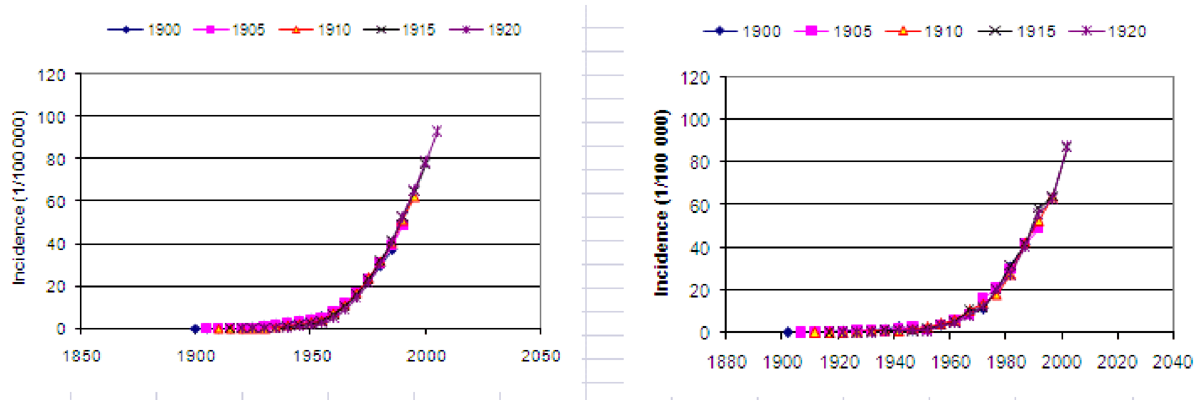


Figure 7. Calculated incidence over calendar time for birth cohorts 1900-1920 is shown to the left and corresponding reported data is given to the right.

Many epidemiologists are used to describe age-specific incidence rates as a function of year of birth. An example is given by Purdue et al. [4] as shown in Figure 8.

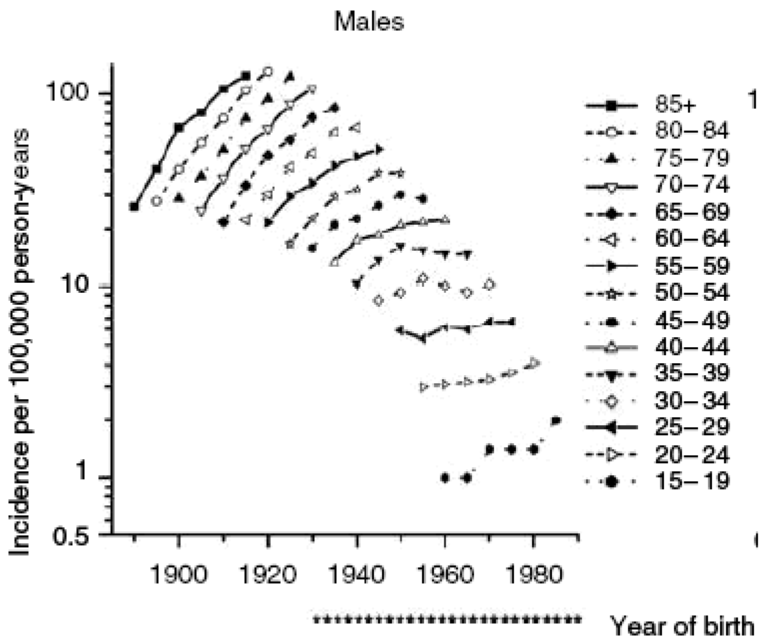


Figure 8. Age-specific melanoma incidence among Caucasians for men stratified by birth cohort year from 1975-1978 through 2000-2004 (from ref. [4])

The physical model used for Figures 4, 5 and 7 can also be used to plot the calculated data as a function of birth cohort year as in Figure 8. The result of this is shown in Figure 9.

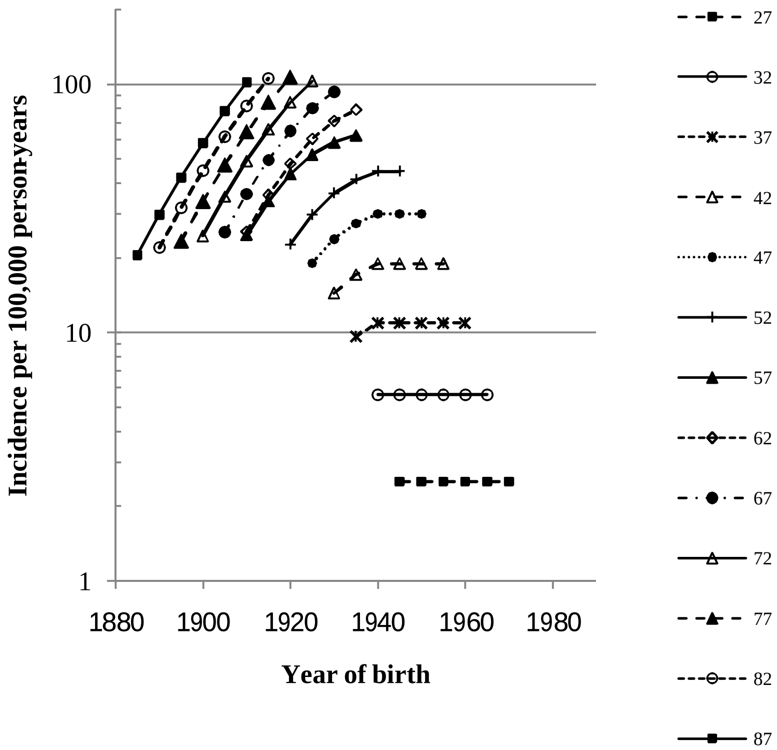


Figure 9. Calculated incidence of melanoma for men in USA according to the model and parameters used in ref. [3].

The sunshine hypothesis

It is easy to use the application to also test the hypothesis that increasing exposure to sunshine and UV radiation is the main explanation to the increasing melanoma trends. On top of the “Calc”-sheet you just define from what year the sun tanning habits started to increase and for how many years it increased (linearly). The total increase is being varied to get the best fit to reported age-standardized data as before but assuming that the natural repair capability is left unchanged. In Figure 10 we have set the start to 1930 increasing up to 1980 (for 50 years).

Define country by "1"		Sun theory	
Sweden	1	Tan started in year	1930
Norway	0	Increase by X times	10,93469812
USA	0	Over Y years	50

Figure 10. To activate the ‘sun-hypothesis’ a passed start year and a duration have to be defined. The total increase of sun-tanning habits is automatically calculated to fit reported age-standardized data.

Again, the calculated age-specific data is ready to be compared with the reported data, see Figures 11 and 6.

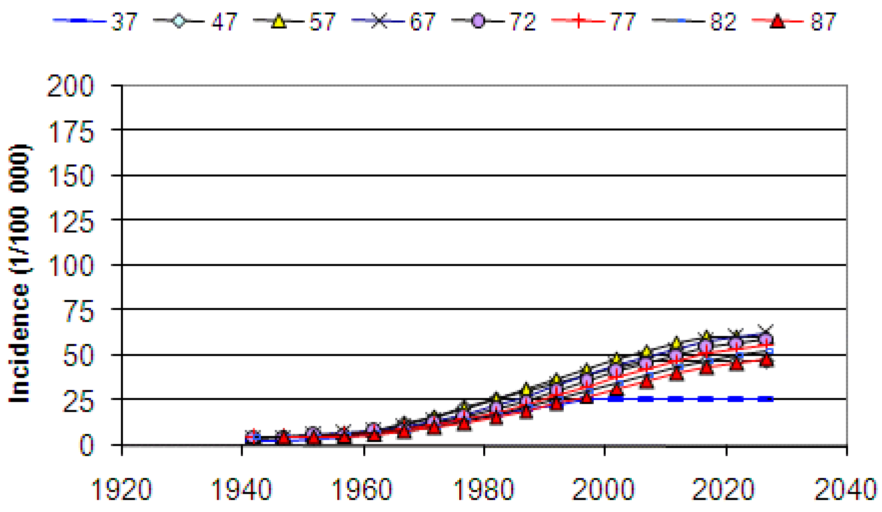


Figure 11. Calculated age-specific incidence for the case that sun-tanning habits have been increasing linearly by 10 times since 1930 up to 1980.

About the application

In its current form the Excel application takes 1.1Mb of memory and is written as a standard 1997-2003 program. If anyone would be interested to use the application for further development and/or use with data from other countries the author suggests to be contacted direct via e-mail to oerjan.hallberg@swipnet.se

References

1. Hallberg Ö. Increasing incidence of malignant melanoma of the skin can be modeled as a response to suddenly imposed environmental stress. Med Sci Monit, 2005; 11(10): CR457-461.
2. Hallberg Ö. A theory and model to explain the skin melanoma epidemic. Melanoma Research, 2006; 16; 115-118. Abstract
3. Hallberg Ö. A reduced repair efficiency can explain increasing melanoma rates. European Journal of Cancer Prevention. 2008;17:147-152. Journal PubMed
4. Purdue MP, Beane Freeman LE, Anderson WF, Tucker MA. Recent trends in incidence of cutaneous melanoma among US Caucasian young adults. J of Invest Derm, epub June 10 2008.