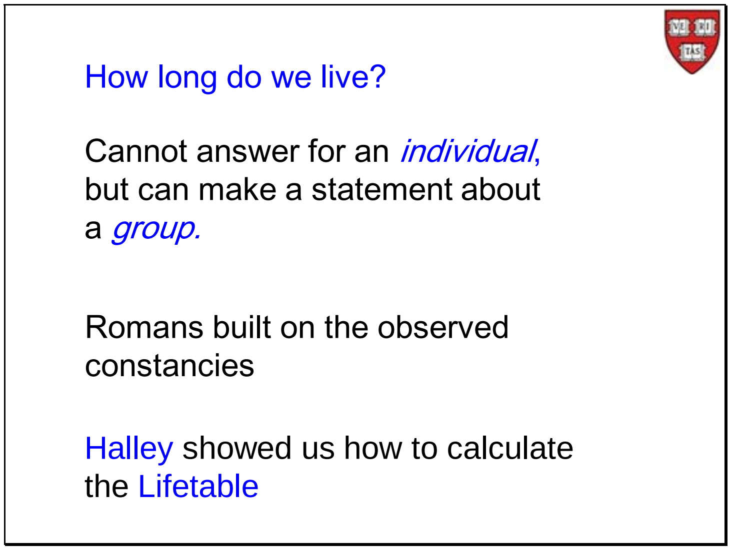
Marcello Pagano

[JOTTER 2 LIFE TABLES]

Material for week 2

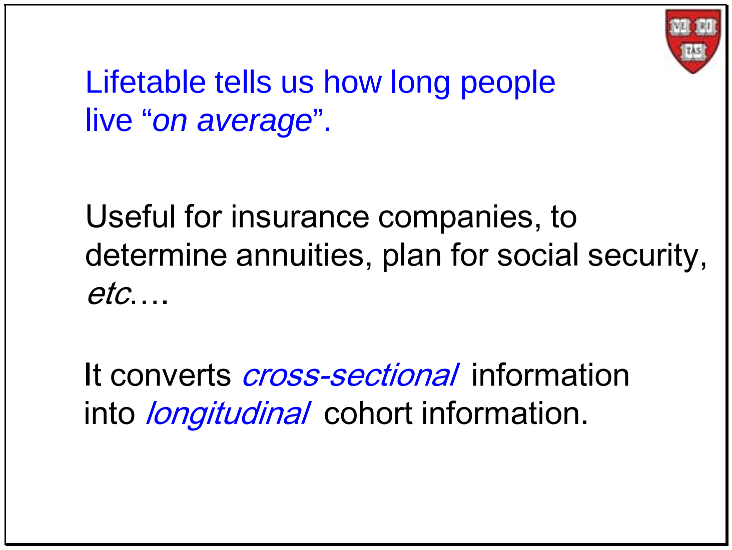


The everlasting question of how long any one of us will live has no answer, of course, but the question is of importance not only to us, but to health policy professionals, to people planning retirement benefits, to insurance companies and to clinical investigators. Currently the best we can do is answer by studying the past, see what happened on average, and project that average onto the future. To enable us to do this in a principled manner, we turn to Halley's method for constructing the life table.

The real answer, of course, is we cannot answer the question, how long will I live? But what we can do is we can answer a question which is, how long do we live? So we can make a statement first about a group as opposed to an individual, and second that group's experience in the past. And that's the answer we provide to the question. How to make such statements so that they have some validity and value is what we now concentrate on.

The Romans observed some constancies that they were able to build on. For example, you need to be able to answer this question if you want to buy annuities, or if you want to buy insurance. How do you answer these questions? At first it was largely based on guesswork. Actually a lot of insurance companies went broke until Halley came along. And he showed us, about 300 years ago, how to calculate what is called the life table.[1] And we've been doing it that way for the last 300 years.

[1] E. Halley, An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives. *Philosophical Transactions of the Royal Society of London* 17 (1693), 596-610 and 654-656.



The life table tells us how long people live on average, and more. It is used for making predictions of what our population will look like in the future.[2] These predictions guide, or, at least, should guide, our policies and plans for the future.

How the table is calculated is very cleverly done. It converts what we call cross-sectional information into longitudinal cohort information. Cross-sectional information is like taking a snapshot of today whereas longitudinal information like a film, or a video of the situation as it evolves over time.

So the idea is, we would love to track a population cohort for the next 100 years or so as it dies off, but we cannot afford to do that. We neither have the time nor the finances to do that. But we can take a snapshot today, or actually a short movie of two year's duration. Technically, we say we've got cross-sectional information and what we're really interested in answering this question is longitudinal or cohort information.

[2] http://www.socialsecurity.gov/oact/NOTES/pdf_studies/study120.pdf

Average Life Span; or how long do we live?

Suppose we have a population of 10 people and we follow them till they die and here are their life spans:

1, 2, 10, 20, 35, 45, 50, 60, 70, 80

So the average life span is:

$$\frac{1+2+10+20+35+45+50+60+70+80}{10} = \boxed{37.3}$$

Let us start with an example with a population of ten individuals and calculate how long they live, on average. Here is how long each one of them lived, for some unit of time: 1,2,10,20,35,45,50,60,70 and 80. We can find their mean, which is 37.3.

What I want to do is I want to put some physical meaning into this and try a little demonstration for you in interpreting this calculation.

[VIDEO PLAYBACK]

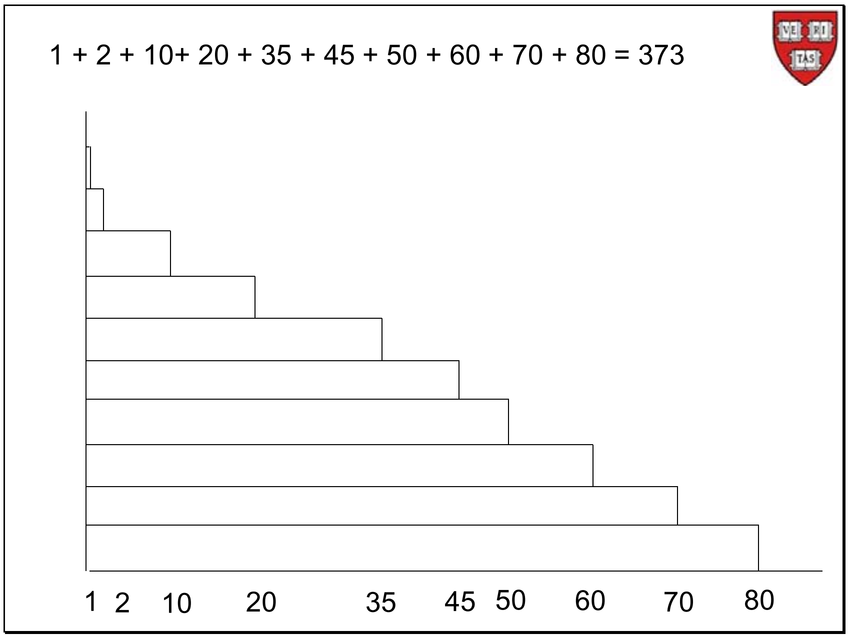
Here we have some sticks. I've cut 10 of them, each stick represents one person, and the length of the stick is how long that person survived. So for example, this one is 80, because the longest was 80. And then there's the 70, the person who survived 70. All the way down to the two and the one. That little one there is a person, a little baby, only survived one unit of time.

Now I want the average survival. So the mean survival should be the sum of all these lengths divided by 10. So one way to do this is I can lay these sticks out. So for example, there's the 80. So that's 80 units. And the next one, I put it in front. And so together they're 150 units. And then the next one, which was six, brings it up to 210 units. And so on, I string them all up like this, make sure I have all 10 of them. And then, what I can do is look at the final length here. And here's the little one. And then divide that by 10. And I've got 373 divided by 10. And so I've got that the average is 37.3. [Error in the film.]

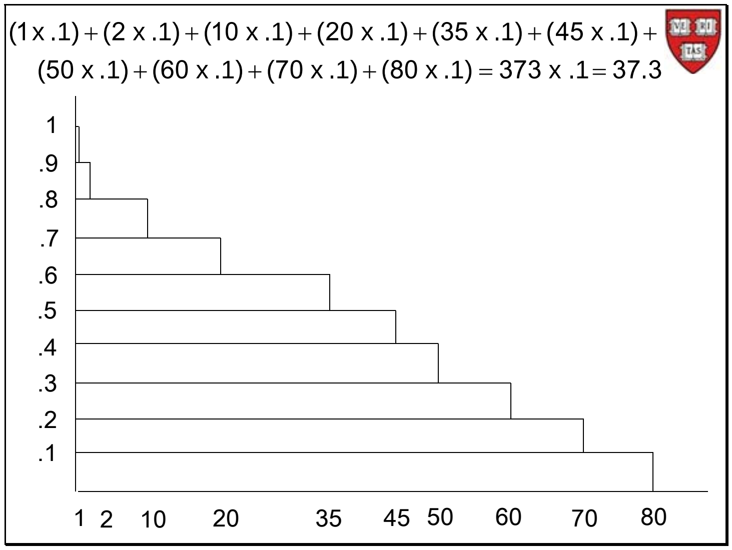
Alternatively, what I could do is rearrange these sticks in a different direction. I can put the 80 at the bottom of the pile like this. And then I can, on top of that, put the 70. And then on top of that, put the 60. And on top of that, the 50. And then the 45. And then the 35. And so on.

And now what I've got, I've got the same sticks. So the total length should be the same. But I've got them in a different configuration. So nothing's changed except the configuration. Let's leave the sticks and actually draw what happens before we explain anything further.

[END VIDEO PLAYBACK]

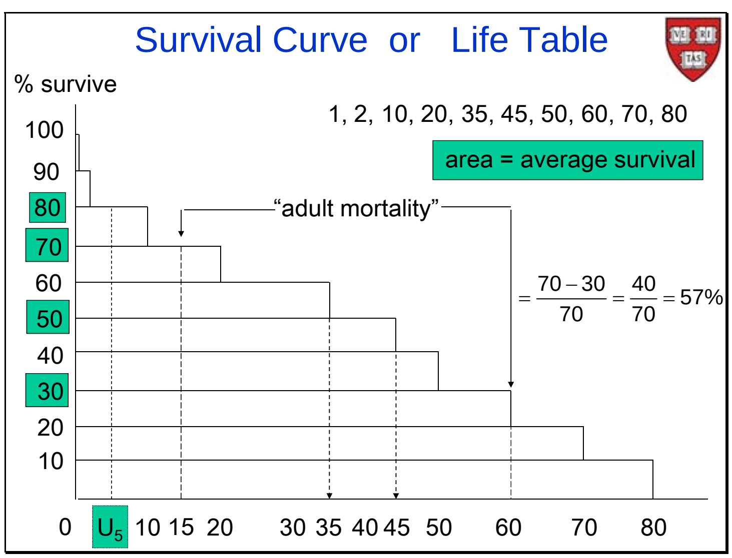


So here is an idealization of those sticks. Here, at the bottom, we have the stick that lasted 80 units. The next one up is 70 units, and so on. So here they are all 10 of them. And as we saw, if we put them lengthwise, it adds up to 373. But now what I want to do is I want to stack them like we did and here is the idealization. If I want to calculate the mean, we need to divide the 373 by 10.



Another way of thinking about this is is to do the division before the summation. Now division by 10 is like multiplying by 0.1. Then you have to add two times 0.1, then add 10 times 0.1, and 20 times 0.1 and 35 times 0.1, et cetera. And that should give you 37.3. It's just simple arithmetic.

Consider what happens if we make the width of the stick 0.1. If we make the width of the stick 0.1, what is this quantity here, 80 times 0.1? 80 times 0.1 is just the area in the bottom rectangle. What about 70 times 0.1? 70 times 0.1 is the area under the second rectangle from the bottom. And so on. So that if we look at this curve, which starts off at one, because we had 10 sticks, each of width 0.1, and follow it all the way down, we get that the mean is just the area under that curve. This curve is called the survival curve or the life table.



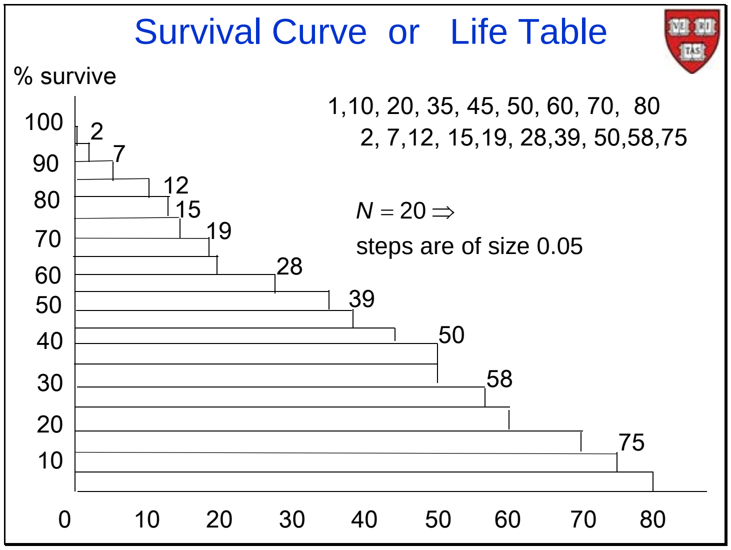
Change the X scale to have them equispaced so that they look more pretty and change the vertical scale to be percentages, so instead of one we will have 100%, 90%, 80%, et cetera. Then we can make the statement that the area is equal to the average survival. So the area under the survival curve is equal to the average survival. The units of the mean is the same as the units of the Xs.

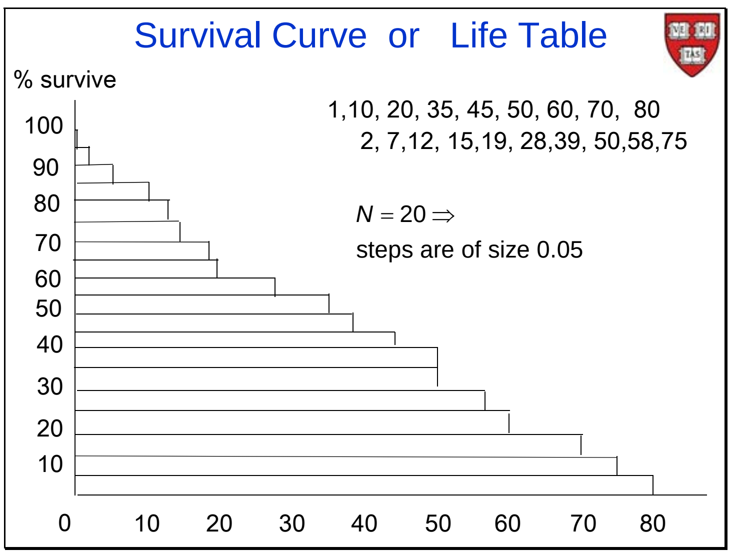
We can also read off other quantities from this curve. For example, if we go up to 50%, on the vertical axis, then we can read off the median survival. Here, in this example, the median survival is going to be anywhere along the horizontal line at 50% which translates to 35 to 45. It's a flat spot, so any number between 35 and 45 will satisfy the definition of the median. Remember the definition of the median: at least 50% above, and at least 50% below. By convention, we can take the midpoint and the midpoint is 40. And that fits in well with the mean that is 37.3.

If this were a survival curve for a large group, such as a nation, then another summary that is often used is the under-five mortality. The way we get that is we look at age five on the horizontal axis and see what that corresponds to on the vertical axis. Here we see that 80% survive to age five. In other words, 20% do not. So the under five mortality in this example, is 20%. Very often you'll find this in the literature to summarize the health experience for a country.

Another summary statistic is the adult survival. The adult mortality, or adult survival, is asking the question, what happens between the ages of 15 and 60? So at 15, there are 70 people who make 15 and survive almost to 60, is 30%. So the adult mortality is defined as 70 minus 30—t hat is how many die between 15 and 60—and divide by the 70 that originally were there. And that gives us 57%. So with this set of data, the adult mortality is 57%.

So those are three ways of gleaning information from this survival curve, summary numbers associated with the survival curve.





Now what happens if instead of 10 people, we have 20 people. So suppose we got this extra data. And we got a new person at 2, 7, 12, 15 and so on. Then you can see what happens. Now that we have 20 the step sizes become 1/2 of what they were before. So the survival curve now looks something like this.

And you can imagine that the more and more and more people we get, the smoother and smoother this survival curve becomes. And we'll end up with something like this:

So as the group size gets bigger the steps get smaller and smaller, eventually leading to a “smooth” curve.

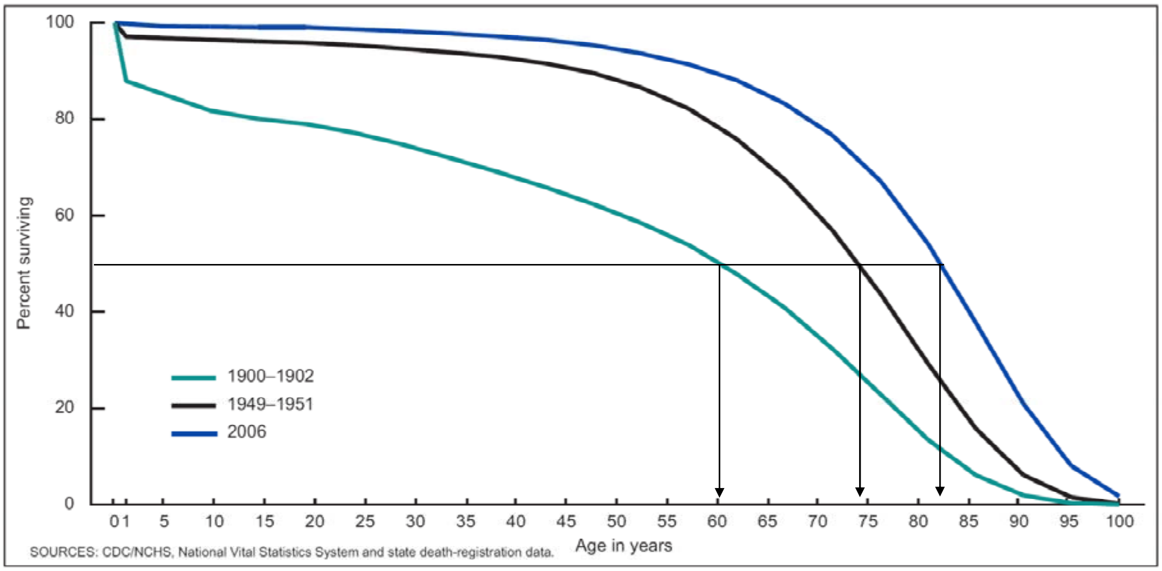


Figure 3. Percentage surviving, by age: Death-registration states, 1900–1902, and United States, 1949–1951 and 2006

So for example, here are three survival curves. And they all refer to the USA. The bottom one is the survival curve from 1900 to 1902. The middle one is the middle of the century, and it refers to the period 1949 to 1951. And the top one refers to 2006.

So here we go across the 20th century. And you can see what has happened to our survival experience. The first thing we notice is that the huge infant mortality, between the ages of zero and one, is slowly disappearing from the turn of the last century, to its middle to the beginning of this century. So here it was at the turn of the century. It's almost gone away. What that has done is it's sort of brought this curve up so that when we look at the average, remember the average is the area under the curve, in the first 50 years of the century, we can see how much we added to the average. In the second half of the century, we added even some more but not as much as we did in the first half of the century.

Another way of looking of looking at the average is to look at the median. As we cut across thye three curves we see that at the beginning of the century it was 60, went up in the middle of the century to about 74 or so, and now it is at about 82, or 83. So we can see this progressive improvement as summarized by the median.

Now the question is how did we construct these curves? These particular ones were constructed by the National Center for Health Statistics. Did they wait 100 years, or so to construct them? Clearly not.

Constructing a life table

To learn how these curves were constructed, we have to go back in time to Halley, Edmund Halley, he of comet fame.

Edmund Halley
1656-1743

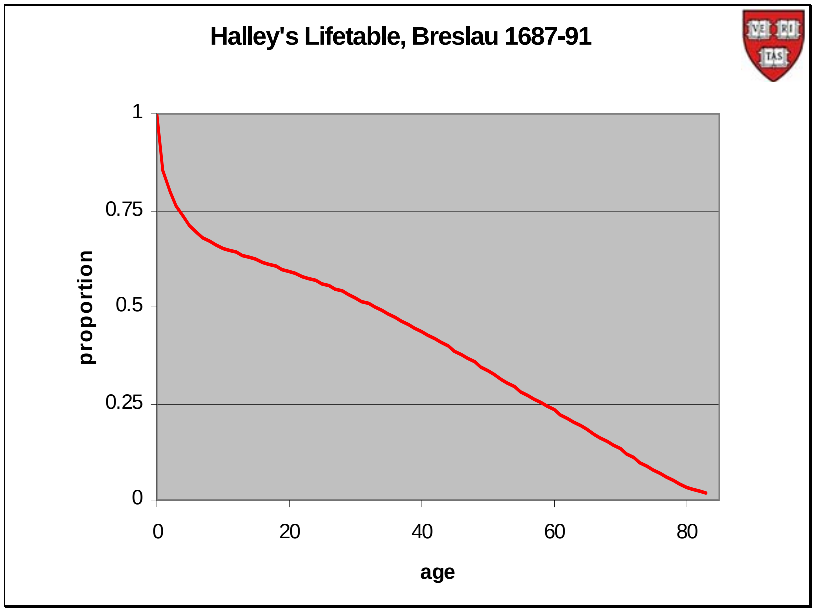
Caspar Neumann
1648-1715

Edmund Halley wrote two papers in statistical methods, or probability, however you want to label them, at the end of the 17th century. And these papers presented his approach to tackling the problem of how to construct the life table. And it just so happened, that the reason why he was approached with the problem, is that someone, by the name of Caspar Neumann, had been collecting some fascinating data.

Neumann was way ahead of his time. What he had done, was collect five years of mortality data from the town of Breslau, which is where he lived, which I think at that time was part of Germany. Now it's part of Poland. He was way ahead of his time, first by collecting data on people, and second, in testing a statistical hypothesis.

He was a religious minister and he was very much interested in astrology. Apparently there is some theory in astrology that says the alignment of the planets when you are born is the same as when you die, or some such. I don't know precisely as I am ignorant of astrology. Anyway, he wanted to test this astrological theory, so he went about collecting data, to test the theory.

He tested it, showed that astrology was nonsense, and at the end of his investigation remained in possession of some wonderful data; five years' worth of it. So a member of the Royal Society sent these data to Halley, who constructed the very first life table using methods that we still use to this day.



And here is Halley's life table of Breslau at the end of the 17th century. And you can see that initial mortality, if one is to trust these data, that initial mortality is extremely high. Only about 66%, of those born reach age 10. So it is a big and precipitous drop. After that, the drop off is almost like a straight line. That's the life table.

Age. Curt.	Per-sons.	Age. Curt.	Per-sons.	Age. Curt.	Per-sons.	Age. Curt.	Per-sons.	Age. Curt.	Per-sons.	Age. Curt.	Per-sons.
1	1000	8	680	15	628	22	586	29	539	36	481
2	855	9	670	16	622	23	579	30	531	37	472
3	798	10	661	17	616	24	573	31	523	38	463
4	760	11	653	18	610	25	567	32	515	39	454
5	732	12	646	19	604	26	560	33	507	40	445
6	710	13	640	20	598	27	553	34	499	41	436
7	692	14	634	21	592	28	546	35	490	42	427
Age. Curt.	Per-sons.	Age. Curt.	Per-sons.	Age. Curt.	Per-sons.	Age. Curt.	Per-sons.	Age. Curt.	Per-sons.	Age. Curt.	Per-sons.
43	417	50	346	57	272	64	202	71	131	78	58
44	407	51	335	58	262	65	192	72	120	79	49
45	397	52	324	59	252	66	182	73	109	80	41
46	387	53	313	60	242	67	172	74	98	81	34
47	377	54	302	61	232	68	162	75	88	82	28
48	367	55	292	62	222	69	152	76	78	83	23
49	357	56	282	63	212	70	142	77	68	84	20

Halley's Lifetable 1693

PHILOSOPHICAL TRANSACTIONS:
Giving some
ACCOUNT
OF THE
Present Undertakings, Studies and Labours
OF THE
INGENIOUS,
In many
Considerable Parts of the WORLD.

VOL. XVII. For the Year 1693.

LONDON:
Printed for S. Smith and B. Walford, Printers to the Royal Society, at the Prince's Arms, in St. Paul's Church-yard. 1694.

Here it is in tabular form. In this form it is easier to explain what Halley did. He started off with 1,000 people. Now take the columns in pairs, and work your way down a pair of columns. So, for example, with the first two columns you start off with 1,000 people in their first year of life. How many will reach into their second year of life? He calculated that number to be 855. Of those, how many reach their third year of life? He calculated 798 would. And of those, 760 will reach the fourth year of life. And so on. Next we read the next pair of columns starting at age 8 and so on, working our way down the columns and then to the right. Eventually we see that of the original 1,000 people born, 20 reach their 84^{th} year.

What he noticed, and what was really ingenious about this, is that he did it step by step. First we can ask, how did he go from 1,000 to 855? What he did was he observed what proportion of all the kids who in those five years died in their first year of life to get the mortality rate in the first year of life. That is the mortality rate he used to get the 855. Then he looked at the proportion of all kids who reached their first birthday but did not reach their second birthday. He then applied that proportion to the 855 to get the number 798. And then what proportion of those who reached their second birthday, reached their third birthday. And so on.

How are these curves constructed? Did we wait 100 years?

Concentrate on the 2006 curve between the ages of 65 and 70.

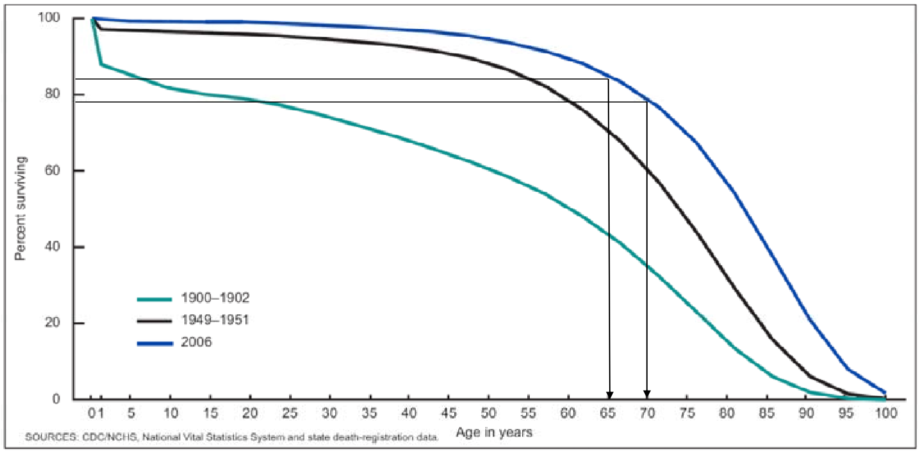


Figure 3. Percentage surviving, by age: Death-registration states, 1900–1902, and United States, 1949–1951 and 2006

We can apply the Halley logic to the modern tables. Focus on the 2006 curve, and ask what happens between the ages of 65 and 70, for example? We first concentrate on the people who have survived to age 65. Suppose we know that point on the curve. Then how do we get the point above age 70? Well if you just tell me what percentage of 65-year-olds will die before they reach 70, and that will give me the point on the curve above the age 70. Then I can repeat that logic to get to the next point, say above 71, and so on. Like that, we can build up the whole curve.

Life Table

We don't know what will happen to the 65 year-olds in 2007 + 65 = 2072 ,
BUT we do know what happened to the 65 year-olds in 2007, and the 66 year-olds in 2007 and the
67 year-olds in 2007 etc…..

The life table methodology constructs a life-experience for a cohort subjected to current mortality rates as it progresses through life, as if the current rates do not change.

Indeed, in 2007 we saw:

That is precisely what Halley did. He said to himself, we don't know what will happen to the 65-year-olds in 2072 because we're not there yet, and we cannot follow this cohort until 2072. But we do know what happened to 65 year olds in 2007. We'll just follow them up for a year or two and we'll see how many die in their subsequent year of life. Similarly I can also tell you what happened to 66 year olds in 2007. I can also tell you about 67 year olds in 2007, and so on.

The next step is to ask if we can project those age-specific mortality rates into the future. One way would be to assume they remain constant into the future. And that is exactly what Halley did.

So the life table methodology constructs a life experience for a cohort. It's a fictional cohort. But it's a cohort subjected to current mortality rates, as that fictional cohort progresses through life. As if the current rates don't change. It is a construct.

Life Table for 2007 USA Age	Probability of dying between ages x to $x + 1$ q_x	Number surviving to age x l_x	Number dying between ages x to $x + 1$ d_x
0–1	0.006761	100,000	676
1–2	0.000460	99,324	46
2–3	0.000286	99,278	28
3–4	0.000218	99,250	22
4–5	0.000176	99,228	17
5–6	0.000164	99,211	16
6–7	0.000151	99,194	15
7–8	0.000140	99,179	14
8–9	0.000124	99,166	12
9–10	0.000105	99,153	10
10–11	0.000091	99,143	9
11–12	0.000094	99,134	9
12–13	0.000132	99,125	13
13–14	0.000209	99,112	21
14–15	0.000314	99,091	31
15–16	0.000426	99,060	42
16–17	0.000529	99,018	52
17–18	0.000627	98,965	62
18–19	0.000715	98,903	71
19–20	0.000796	98,832	79

So the second column in this table for 2007 tells us what age-specific mortality rates were actually observed from the one-two year followup starting in 2007 for the first 20 years of life. That's because I can only put 20 years of life on this slide.

If, in 2007, I look at all the babies born in 2007 and I follow them up for one year—so that means for babies born on January 1, 2007, I have to follow up until January 1, 2008, for babies born on January 2, I have to follow to January 2, 2008, etc.. So every baby is followed for a year. That means it takes me two years to collect my data. Because for the babies born on December 31, I have to follow them up for two years. So for all babies born in 2007 the proportion who passed away within one year of being born is 0.006761.

Now, what about kids who were born in 2006, and thus turn two in 2007.in their second year of life in 2007? What proportion of them died within a year of their first birthday? And the answer is, 0.000460. So, once again I need only follow them up for two years. So for the ones who turned one on January 1, 2007, I have to follow up for a year to 2008. All the way down to December, I have to follow up for a year to do that. So within two years, I get both of these points. And the same is true for those between the ages of two and three. So those in their third year of life, I find by following them up for two years, that I get this number, which is 0.000286. So that's the proportion who die.

So in fact, this whole column here, this whole first column here, all these numbers and all the way down to 100, or whatever the lower number is, is based on observing this cohort over this very, very short amount of time. Okay, so these are the actual age-specific mortality rates for the 2007 year. Not for the 2007 cohort, but for the 2007 year.

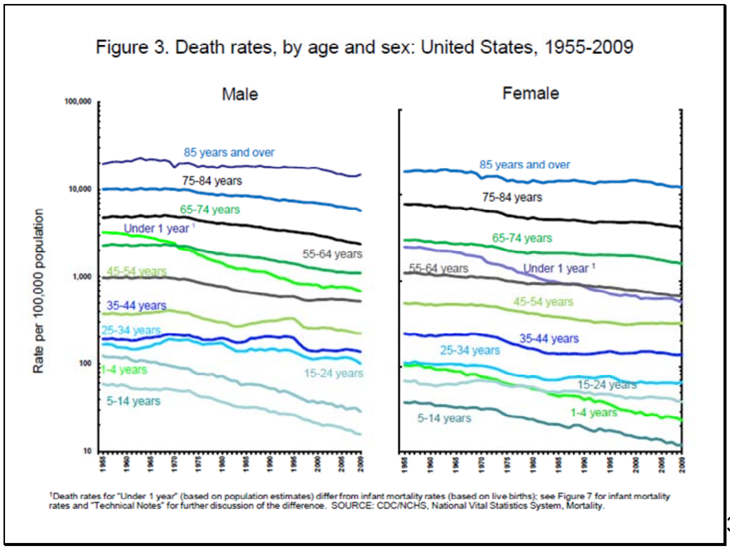
Now, to actually ask that other question, which is, all right, so after two years I know what happened to babies in their first year of life. What about the babies born in 2007 in their second year of life? Now I have to move into 2008, 2009 to get that. In the third year of life, I have to move into 2009, 2010. And so on, if I'm going to follow this cohort through life. And I can't. Very quickly, I run out of time because we're in 2012. All right? But, what I can do is I can apply these numbers I've got here to that cohort.

So let me construct a fictional cohort and subject it to the current mortality rates. So let me start with 100,000. Apply this proportion dying. So of those 100,000, we're saying, 0.006761 are going to die. So that gives me 676. So 676 are going to die in their first year of life. What does that leave with me with? Well 100,000 minus 676 which is 99,324. Let me look at this proportion of those that are going to die. So the 0.000460 times this will tell me how many of those kids will die in the following year. And that gives me 46. Subtract 46 from 99,324 gives me 99,278.

How many of those are going to die in the subsequent year? I multiply by 0.000286 and I get that 28. And subtract that from 278 and I get 99,250. And so on. And I keep on doing this.

So I'm creating a fictional cohort and applying today's age-specific mortalities to that. And what do I end up with here and carrying on down is my survival curve for this cohort of 100,000 people.

So remember, our survival curves start off with one. So to get that I just divide by 100,000 people. They introduced the 100,000 just to make the arithmetic look simpler. But just once you're done, just divide by 100,000 and then that's our survival curve.



[3]

The question then is, how good is this survival curve? Well we won't know for a hundred years or more, of course, but if we go back in time, we see that the age specific mortality rates are pretty constant. Most of them are going down. If there is a trend, that trend is going down. There are some exceptions such as the Vietnam War for males. If the trends continue to go down, then that will mean that our estimates of the survival curve are conservative and that people will actually live longer than predicted.

---

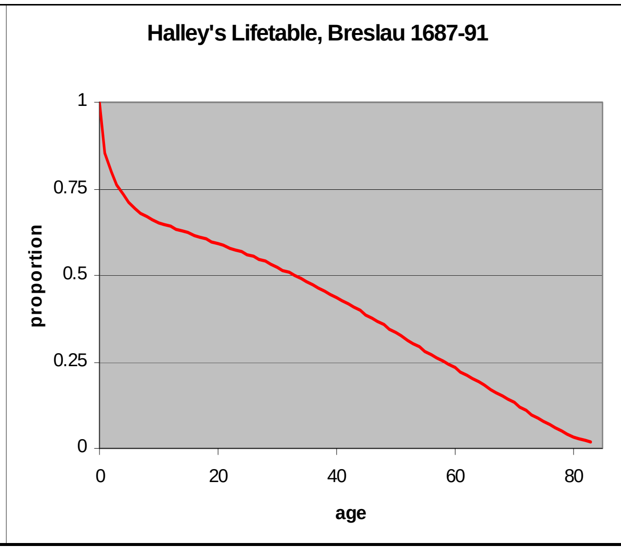
[3] http://www.ssa.gov/oact/NOTES/as120/LifeTables_Body.html#wp1176553

Stable Population Assumption

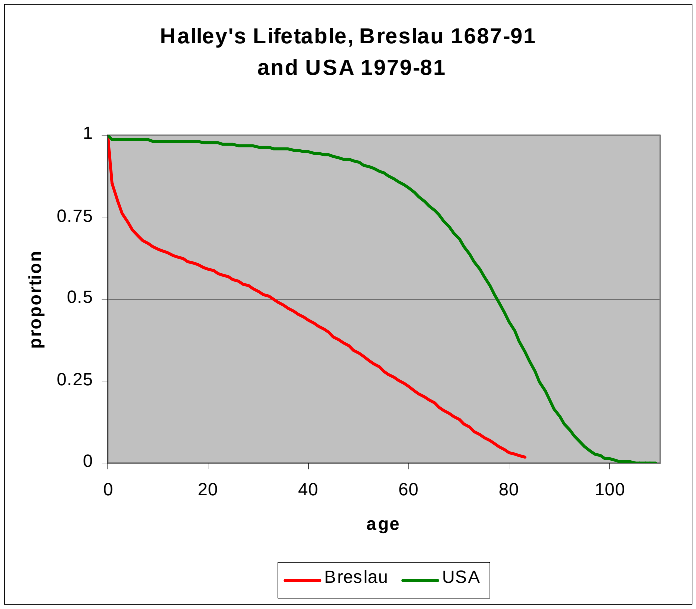
Halley did not have a population count for Breslau, so he made the extra assumption that the population was stable over the five years that Neumann counted deaths meaning that he assumed that the number of deaths equaled the number of births.

He actually then went on to infer the population size of Breslau during that period, from his lifetable!

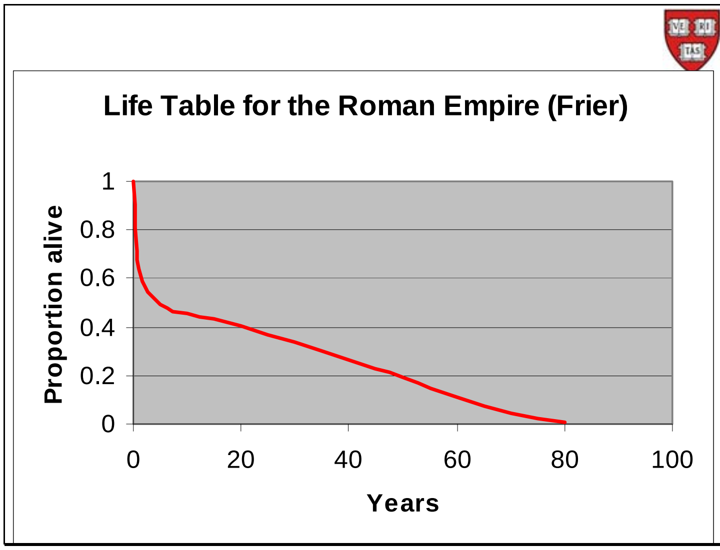
I should add that Halley made one extra assumption. For the 2007 life table we started from the proportions dying within age groups. To get those we had the number dying in the numerator and the number alive at the beginning in the denominator. Halley did not have a census so he did not know how many were alive, he had to estimate how many people were in Breslau. So he made to make an extra assumption, which was the stable population assumption, which basically says the number of kids born is equal to the number of kids dying within any one year. He recognized that he was a little bit off in his calculations because of that assumption. What we inherited from his work is a method for doing the calculations.



And I repeat, it has proved invaluable in that we still use it more than 300 years later.



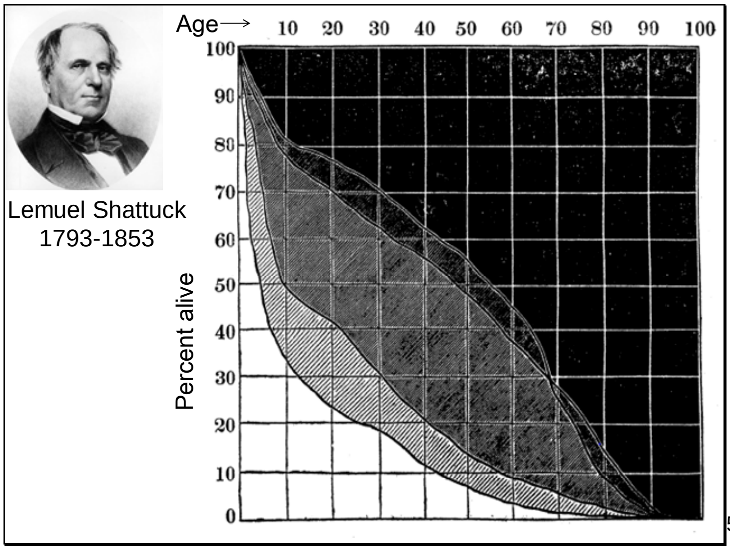
Here is Halley’s life table of Breslau for the period 1687—91; it is the red line. We have superimposed the US life table for 1979—1981. We can see a few differences of what's happened. The biggest, of course, is what happens in the early years of life. What used to happen 300 years ago was much worse than what happens nowadays. Also note that now we sort of coast along until what, age 50 or so. And then we start dying at a higher rate. Whereas in Breslau's time, it was almost a constant rate from age 10 or so.



[4]

Here is what Frier thinks that the life table at the time of the Roman Empire looks like. And this is based on looking at tombs and records and seeing how old people were when they died and a very imaginative reconstruction of some data. If one is to believe these data, then we see a big precipice the first 5, 10 years of life and we are almost down to 40% by age ten.

[4] http://www.richardcarrier.info/lifetbl.html : Landlords and Tenants in Imperial Rome, Bruce W. Frier, Princeton University Press, 1980.



[5]

This one is my favorite. This is out of a report by Lemuel Shattuck. Lemuel Shattuck was the chair of a committee that did a survey of Boston in the middle of the 19th century, and he apparently wrote most of the report. He included only one graphic in all of it and it is this one.

He was making a case for improving public health in Boston. He attacked the quality of the air, the filth in the abodes, and the general poor level of hygiene in the city. Shattuck is known as the father of public health in this country. He outlined what a state department of public health should look like and, in fact, Massachusetts was the first state public health department and it based on his blueprint. Unfortunately, it happened after he died so he didn't get the credit that he deserved.

This graph shows four separate life tables. He chose a city in England that had roughly the same level of industrialization as Boston, Preston, and over roughly the same period of time. That is the top curve here. Be careful because here it crosses the next curve. The next curve down is Newton, Massachusetts, a rural town at the time. You see that Newton and Preston are very similar and as I said, by age 70, they crossed over.

Now the surprise is that just below that we get the life table for Boston. We see that, in Boston, about 50% of all kids died by age 10. And this he was attributing to the pollution and the living conditions and sanitation, et cetera, for the inhabitants of Boston.

---

[5] Shattuck picture: http://www.mass.gov/eohhs/consumer/physical-health-treatment/health-care-facilities/public-health-hospitals/shattuck/mission-and-history.html

He also makes the point that the poorer people had it even worse. The bottom curve is still Boston, but it represents the Catholics in Boston. That life table goes down to 30% by age ten. So by age 10, 70% of all kids died.

So a very powerful picture that actually did move the legislature into action.

Dean Briggs of Harvard says, “The peculiar evil in cigarettes I leave for scientific men to explain; I know merely that among college students the excessive cigarette smokers are recognized even by other smokers as representing the feeblest form of intellectual and moral life.”

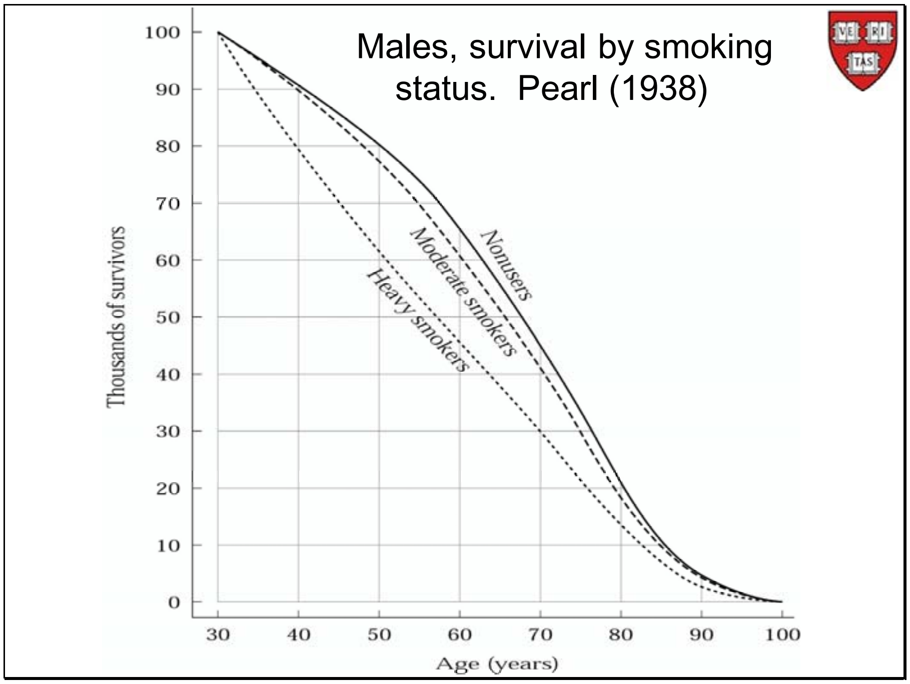
Healthy Living, Book Two
C.A. Winslow 1920 p.185

Another example, this time having to do with smoking. Here is what the dean at Harvard, Dean Briggs, had to say about smoking. And this was reported in 1920, and it is surprising that they recognized the evil in cigarettes as far back as then:

"I leave for scientific men to explain. I know merely that among college students the excessive cigarette smokers are recognized even by other smokers as representing the feeblest form of intellectual and moral life."

It was not until 1939 that Muller came out with his famous, but largely ignored, study that showed that smoking cause lung cancer.[6]

[6] Muller FH. Tabakmissbrauch und Lungencarcinom. Z Krebsforsch 1939;49:57-85.



Around that time, Ray Pearl published an article in 1938, of a study he carried out where he divided a cohort of men into three categories: “nonusers”, moderate smokers, and heavy smokers. Above are the survival curves derived from these three groups.

He started the table at age 30 primarily because he was interested in the effects of smoking, so he concentrated on adults. What we see is a much steeper drop for the heavy smokers when compared to the other two groups. The differences between the moderate smokers and the nonusers is there, but not as pronounced. If we look at the median survival, we see that the heavy smokers have a median age of about 57, or so, and the nonusers have a median age of about 67. So roughly speaking, just by looking at medians, one loses about 10 years of life on average because of membership in the heavy smoking group.

There is one other interesting observation to be made, and that is that death, like taxes, is inevitable. It happen to all of us. So what he drew here is that by age 100, everybody is dead. OK. But look at what's happening here. If we look at say, age 70, there's only 30% of heavy smokers left. So if we started with 100 heavy smokers, there are only 30 left by age 70, whereas for the nonusers, there are about what, 45% left at age 70. That means there would be 45 left of the original 100.

Now here's the interesting thing. By age 100, they're all going to die. So in the next 30 years, between 70 and 100, these 30 heavy smokers will die, whereas in the same amount of time 45 nonusers die. That tells us that if you have the same amount of time to see 30 die as 45, then the 45 must die at a faster rate. That says to us that the mortality rate for the nonusers is going to be higher, than it is for the smokers.

This is an example of what has been called the healthy worker effect. [7] Summarised by Pearl as, "This is presumably an expression of the residual effect of the heavily selective character of the mortality in the earlier years of the groups damaged by the agent (in this case tobacco). On this view those individuals in the damaged groups who survive to 70 or thereabouts are such tough and resistant specimens that thereafter tobacco does them no further measurable harm as a group." [Science, March 4 1938]

TABLE 1

THE DEATH RATE (1,000 q_x) AND SURVIVORSHIP (l_x) FUNCTIONS, AT FIVE-YEAR INTERVALS, STARTING AT AGE 30, OF (a) NON-USERS OF TOBACCO; (b) MODERATE SMOKERS WHO DID NOT CHEW TOBACCO OR TAKE SNUFF; (c) HEAVY SMOKERS WHO DID NOT CHEW TOBACCO OR TAKE SNUFF. WHITE MALES

Age	Non-users		Moderate smokers		Heavy smokers	
	1,000 q_x	l_x	1,000 q_x	l_x	1,000 q_x	l_x
30	8.18	100,000	7.86	100,000	16.89	100,000
35	8.78	95,883	9.63	95,804	21.27	90,943
40	10.01	91,546	11.89	90,883	23.91	81,191
45	12.04	86,730	14.80	85,129	25.69	71,665
50	15.16	81,160	18.61	78,436	27.49	62,699
55	19.82	74,538	23.67	70,712	30.09	54,277
60	26.73	66.564	30.49	61,911	34.29	46,226
65	36.88	57,018	39.83	52,082	41.20	38,328
70	51.69	45,919	52.84	41,431	52.72	30,393
75	73.02	33,767	71.28	30,455	72.33	22,338
80	103.22	21,737	97.95	19,945	100.44	14,494
85	142.78	11,597	136.50	10,987	139.48	7,865
90	197.49	4,753	190.23	4,686	193.68	3,292
95	273.2	1,320	265.1	1,366	268.9	938

Here are the actual numbers used by Pearl. So if you look, by age 70, of the non-smokers, there are 45,919. Of the moderate smokers, there's 41,431. And of the heavy smokers, there's 30,393. The mortality rates are roughly the same at age 70; for the three groups they are, 51.69, 52.84 and 52.72 per 1,000, respectively. At age 75, the mortality rates for the nonusers is 73, which is higher than for the heavy smokers, which is 72. For the nonusers at age 80, it is 103, whereas it is 100 for the heavy smokers. At

[7] **Summary of Several Male Life Tables,** William Ogle *Journal of the Royal Statistical Society*, Vol. 50, No. 4. (Dec., 1887), pp. 648-652.

age 85, it is 143 versus 139, and at age 90 it is 197 versus 193. Finally at age 95 it is 273 versus 269 per 1,000.

Do not, by any means read this to mean that you should not smoke until age 70 and then start smoking.

You will see in your epidemiology studies that if the outcome determines, or impacts, your membership in a group, then it makes subsequent conclusions suspect.

Now the other interesting thing is that his conclusion is that smoking is not very good for you. But what struck me when I first read this was that he published this in 1938. And he published it in a very reputable journal, Science. But somehow we ignored this message. It's all to the power of the tobacco industry that they were able to, because of advertising, and other means have such an impact on us.

Healthy worker effect

Here, just as is usually the case in our experience in studies of this sort, the differences between the usage groups in specific mortality rates, as indicated by q_x , practically disappear from about age 70 on. This is presumably an expression of the residual effect of the heavily selective character of the mortality in the earlier years in the groups damaged by the agent (in this case tobacco). On this view those individuals in the damaged groups who survive to 70 or thereabouts are such tough and resistant specimens that thereafter tobacco does them no further measurable harm as a group.

Smoking tobacco effect

However envisaged, the net conclusion is clear. In this sizable material the smoking of tobacco was statistically associated with an impairment of life duration, and the amount or degree of this impairment increased as the habitual amount of smoking increased.

Tobacco Smoking and Longevity
Raymond Pearl
Science, New Series, Volume 87, Issue 2253
(Mar. 4, 1938),216-217.

WHO

http://www.theglobaleducationproject.org/earth/index.php

I'd like to direct you to the World Health Organization. They have a lovely little website that does a lot of these calculations around the world, what life expectancy is around the world. So, go to the website and you'll see the address just above.

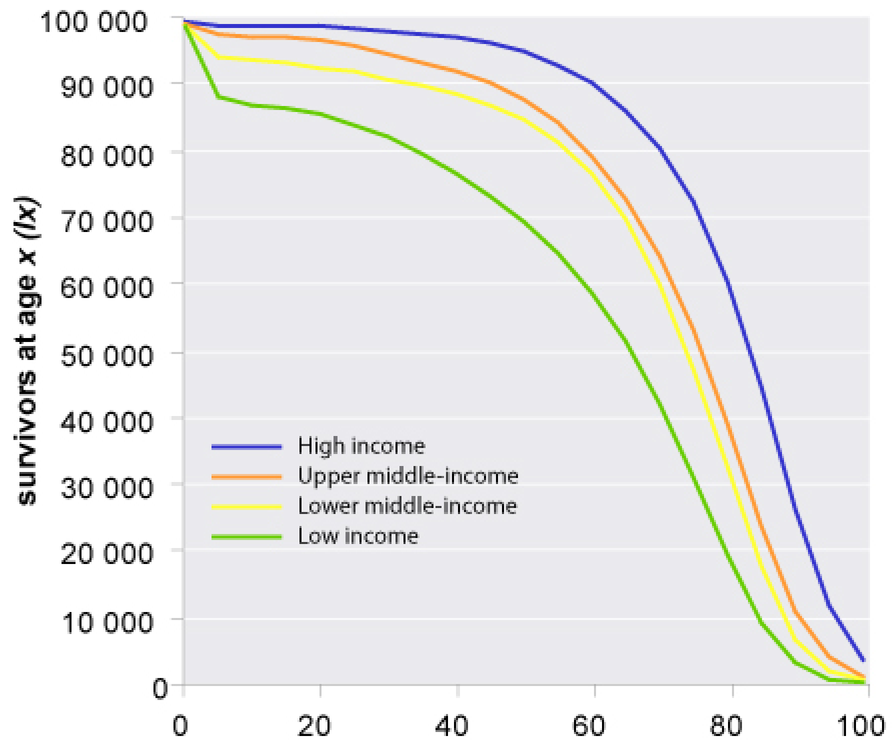
Once you get there, you should be able to view interactive graphics that are full of information.

You see a map of the world that allows you to zoom in on certain parts of the world. If you scroll over the country, you can see various information about the country. So, for example, Brazil-- life expectancy in Brazil is 73. If you go to India, life expectancy is 65, and so on.

And you can rank the countries alphabetically on the left. Or you can rank them by life expectancy. You can see that Malawi has the lowest. Afghanistan, Zambia, et cetera-- the lowest ones. If you want the highest ones, you just click again. We see that the highest is in Japan, and San Marino Republic in Italy, et cetera.

So you can have fun with this. And it's very instructive. And you can see how things vary around the world.

So it's worth a few minutes of your time to learn a lot about how life expectancy is distributed around the world.



http://www.who.int/gho/mortality_burden_disease/life_tables/life_tables/en/index.html

One more site for you to visit to study some important differences in survival experiences around the world.