

James A Hanley

Department of Epidemiology, Biostatistics and Occupational Health  
McGill University, Montreal, Canada.

possible outlets: Epidemiology, American J. of Epidemiology, ...

.

Comments welcome

## Abstract

The terms force of mortality, incidence density and hazard rate and involve the same concepts. However usage tends to divide along professional lines. Actuaries and epidemiologists tend to see the force of mortality and incidence density respectively as empirical quantities, and deal with variations in these over age or time by segregating them into age- or time-bands. Statisticians usually first define the hazard rate theoretically – using continuous functions and mathematical limits – and then develop estimators for it. The resulting mathematical divide is particularly acute in the case of the “exponential” formula allowing one to calculate the risk over a given age- or time-period from the continuous function reflecting the variation in force of mortality, the incidence density or the hazard rate over this period. Since the only formal derivation until now involves advanced calculus, few epidemiology textbooks present, and fewer still fully explain, this formula. Increased understanding of this link, and its underpinnings, is especially important today, as the familiar Kaplan-Meier estimator of a cumulative incidence proportion or risk is gradually being replaced by the Nelson-Aalen estimator. Likewise, investigators increasingly use both non-parametric and parametric statistical models to calculate profile-specific x-year risks, risk differences, and numbers needed to treat, and to test proportional hazards via log[survival] plots. In part I, we first use a worked example to illustrate the concepts common to the force of mortality, the hazard function, and incidence density functions. We then revisit the 1832 definition of the force of mortality and how a person-year was conceptualized, and use a striking 2010 graph to re-emphasize the centrality of time. In part II, we extend the 1832 conceptualization, and use the probability of a specific realization of a Poisson random variate to de-mystify the formula linking an incidence function and risk.

## 0 Introduction and outline

The terms mortality, hazard rate, incidence (rate), and incidence density all involve the same concepts, but those that involve a mathematical limit (derivative) or integral make many epidemiologists uncomfortable. Indeed, although epidemiologists are comfortable with the concept of full-time equivalents in measuring staff sizes, this comfort level does not always extend to the concept of an intern-month or intern-year, or to converting an incidence function to a cumulative incidence proportion or risk. As a result, epidemiologists may be unsure as to how to turn an injury rate of say 0.095 needle-stick injuries per intern-month into a 12-month cumulative incidence or risk, and of what assumptions are involved. Indeed, few textbooks present, and fewer still explain, the formula linking incidence and risk.

Increased understanding of this link is all the more critical nowadays, as the familiar Kaplan-Meier estimate of risk is gradually being replaced by the Nelson-Aalen one, and as investigators use non-parametric and parametric statistical models to calculate profile-specific x-year risks (Schröder 2009), risk differences, and numbers needed to treat (Ridker, 2008).

The second of this pair of articles addresses our main objective – demystifying (and giving a simpler derivation of) the formula used to convert an incidence function to a cumulative incidence (proportion-type) ‘rate.’ To help with the heuristics, the first of the two parts provides an orientation to incidence density and hazard functions themselves.

We agree that while “the calculation of frequencies of disease is the most basic tool for epidemiology,” the “fundamental concept of the nature of the incidence rate and its calculation in dynamic populations is often not well explained, especially in introductory courses. We applaud the IJE for beginning their Education Corner with an item that goes back to the basics. Since one can never be too clear about these, section 1 reviews the concept of a hazard or an incidence density in a single interval defined by a timepoint  $t$ . We use data from a dynamic population experience to measure the force of mortality (hazard function, incidence density) over the life course, and comment on some 19th century attempts to fit it using smooth parametric functions.

Since the role of time, and time-units, in the measurement of the incidence function is often neglected, section 2 presents a striking time-graph to illustrate how critical time units are, and how easily they are overlooked, or even misunderstood.

Section 3 revisits the first definition of the force of mortality, a term coined by TR Edmonds, an actuary colleague of William Farr. Whereas most of us think of a person-year simply as the unit by which we measure an aggregated amount of experience of different persons, this actuary’s concept of a person-year was somewhat different. His concept leads to a simple and constructive derivation of the link between incidence functions and the cumulative incidence proportion or risk.

## 1 Incidence rate; incidence density; hazard rate; force of mortality

Epidemiology texts tend to define an incidence rate or incidence density using ‘numbers of events’ and person-time denominators, in intervals of time. Statistical texts define their equivalents, hazard rate or force of mortality, as mathematical limits involving probability distribution functions and survival functions measured in continuous time. The twain seldom meet.

The 2009 Wikipedia<sup>1</sup> entry for the ‘instantaneous hazard rate’, presumably prepared by a statistician, is a case in point. It defines it, in words first, as “the limit of number of events per unit time divided by number at risk as time interval decreases”, and then in symbols:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{\text{observed events in interval } [t, t + \Delta t] / N(t)}{\Delta t}.$$

The entry did not define  $N(t)$ , but presumably it was used to denote the number at risk at time  $t$ , and allowed for the possibility that it might be different at time  $t + \Delta t$ , and at intermediate times. Let us set aside for now the focus on the limit, and consider on a finite<sup>2</sup> interval  $[t, t + \Delta t)$ . If we replace  $N(t)$  by  $\bar{N}$ , the average no. persons being observed during the interval  $(t, t + \Delta t)$ ,<sup>3</sup> and re-arrange the expression so that the denominator is

---

<sup>1</sup>The entry has since been edited.

<sup>2</sup>As we show below, if the  $\Delta t$  is reasonably small, the value of the expression will not change appreciably if it is made smaller still.

<sup>3</sup>Note that  $\bar{N} \times \Delta t$  is also the integral of the  $N(t)$  function over the interval in question.

in the form of a ‘persons  $\times$  time’ product, then we get, for this finite interval, the ratio

$$\frac{\text{no. events in } (t, t + \Delta t)}{\bar{N} \times \Delta t} = \frac{\text{no. events in } (t, t + \Delta t)}{\text{Population-Time in } (t, t + \Delta t)}$$

which takes the form of incidence density, a term introduced to epidemiology by Miettinen (1976):

Incidence density (“force of morbidity” or “force of mortality”) – perhaps the most fundamental measure of the occurrence of illness – is the number of new cases divided by the population-time (person-years of observation) in which they occur.

In light of this, is the hazard rate or force of mortality the same as the ‘short-term incidence density’? A concrete example in which we calculate the hazard rate or force of mortality at a given age  $t$ , based on the numbers of deaths, and person-years of observation in the USA population over the period 2000-2006, shows that it is,. Thus, there is no need to be frightened by the mathematical limit, or by having to actually carry out the calculation using a tiny interval, or by the prospect of being trapped in an ‘instant’ or ‘moment’– terms we will use for a time-interval of zero length. We never *consider* the zero length interval that some people equate with the term instantaneous: we simply measure the ratio as we *approach* that zero length interval.

The ‘almost-raw’ data for our calculations, as well as the full ID function

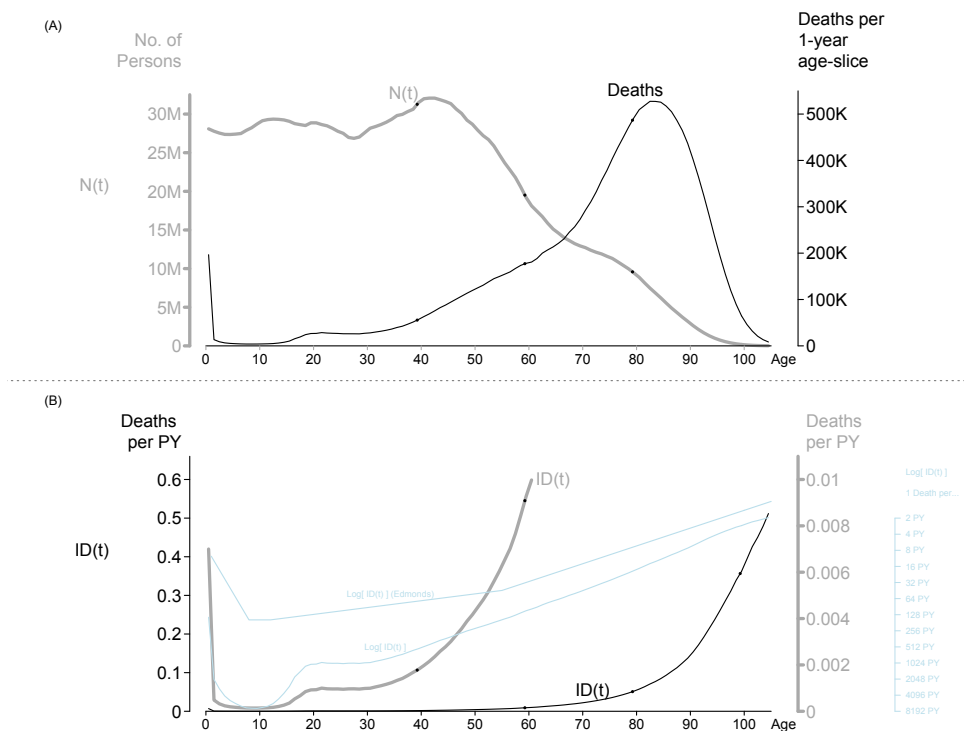


Figure 1: Age structure of, and age-specific numbers of deaths recorded in, the (dynamic) USA population observed over the period January 1, 2000 to December 31, 2006 **(A)**, along with the age-specific death rates derived from them **(B)**. Source: Human Mortality Database, <http://www.mortality.org/>. For each (continuous) value of age ( $t$ ), shown in (A) is  $N(t)$ , the number of persons who were ‘exactly  $t$  years of age’ on *some* date in the 7 calendar years. Thus the numbers of person-years lived in any age-interval is the integral of (area under) the  $N(t)$  curve over the interval in question. Most residents contributed 7 person-years each to the overall total of 2,000 million person-years; some contributed fewer – mostly to either the younger or older end of the person-years distribution. The numbers of deaths for any age-interval is the integral of the ‘deaths per 1-year-of-age time slice’ curve (axis on right) over the interval in question. Shown in (B) are the full (in black, left axis) and ‘below age 60’ portion (grey, right axis) of the  $ID(t)$ , or force of mortality or hazard rate function. The  $ID(t)$  function ranges from a nadir of 0.000014 year<sup>-1</sup> at approx.  $t = 10$ , to 0.51 year<sup>-1</sup> at age  $t = 105$ . The log – to the base 2, so that we can easily measure doubling times – of the  $ID(t)$  function is shown on yet another scale. Gompertz’ Law of Mortality, in which the rate ‘doubling time’ is approximately constant (the logs of the rates are approximately linear) appears to hold true for the age range 30-90. For historical interest, Edmonds’ piecewise-linear log(ID) curve, based on data from early 1800s, is also shown on this scale.

derived from it, are shown in Figure 1. The population sizes and numbers of deaths were only available for 1-year age bins; thus, in order to display them as a continuous function of age, the numbers were smoothed out using a spline function. The integral under the curve over any specific age-interval is the numbers of person-years lived in, or the numbers of deaths recorded for, the age-interval at issue. Over the age span 0-105, there were approximately 17 million deaths in just over 2,000 million person-years.<sup>4</sup>

We take as illustration the ID or force of mortality or hazard rate at the (deliberately selected to be a non-integer) age  $t = 39.25$ . Technically, persons are only exactly 39.25 for a moment (infinitesimal, since a moment has no duration) and so we can only consider the calculation over say the finite interval  $(t - \frac{\Delta t}{2}, t + \frac{\Delta t}{2})$ , of width  $\Delta t$ , that includes  $t = 39.25$ .<sup>5</sup> Table 1 shows the calculations with successively shorter intervals. The IDs would ultimately become unstable if we considered intervals as short as  $39.25y \pm 3$  hours, say or  $79.25y \pm 1$  hour. However, even as we narrow the intervals from a year to a day – or to minutes and seconds and nanoseconds if we ignore sampling variability and restrict our focus to the theoretical (i.e., abstract, expected) values – the ID's are practically unchanged. As Figure 1(A) shows, the  $ID(t)$  function does not change abruptly; it changes continuously – but slowly.<sup>6</sup>

---

<sup>4</sup>Thus, the overall ID was  $0.0085 \text{ year}^{-1}$ ; its reciprocal – 118 years – reflects the fact that this population experience is younger than in the current lifetable (expectation of life at birth: 77 years) calculated from these data.

<sup>5</sup>Since it doesn't fundamentally alter the concept, readers will find it easier, as we do here, to take  $t$  to be the *center*, rather than the *left boundary*, of the interval. The use of successively smaller intervals centered on  $a$  does cause some mathematical difficulties at  $t = 0$ , and explains why the limit is typically approached from the right.

<sup>6</sup>over the 1-year interval centered on  $t = 39.25$ , the ID increases by about 0.021% per

Table 1: Incidence density (ID), calculated for (successively smaller) intervals, of width  $\Delta t$ , centered on 3 different timepoints

$\Delta t$	$t = \text{age } 39.25$			$t = \text{age } 59.25$			$t = \text{age } 79.25$		
	P-T*	Deaths	ID	P-T	Deaths	ID	P-T	Deaths	ID
1 year	. 31.255	55,590	177.9	. 19.520	177,133	907.5	. 9.578	486,785	5,082.3
1 month	. 2.605	4,629	177.7	. 1.626	14,772	908.6	. 0.798	40,567	5,082.5
1 week	. 0.601	1,068	177.7	. 0.375	3,409	908.6	. 0.184	9,362	5,082.5
1 day	. 0.086	152	177.7	. 0.053	486	908.6	. 0.026	1,334	5,082.5

\*P-T Units: 1 million person-years ID Units: deaths / 100,000 years

Based on population-sizes and numbers of deaths, USA 2000-2006.  
 Source: Human Mortality Database, <http://www.mortality.org/>

Thus, the only reasons to be ‘instantaneous’ about it are if one wished to have a continuous smooth curve, especially one with a functional form, to shorten tedious annuity calculations or to compute an x-year risk (cumulative incidence), or to be able to provide an accurate break-even premium for 1-day term insurance for a large group of people.

We suspect that part of the ‘divide’ between statisticians and epidemiologists in this matter has to do with two different – but operationally equivalent – ways they define the hazard and the incidence density. Statisticians tend to first view it as a *theoretical* quantity and define it – in the abstract – as a (conditional) ‘probability per unit time’ for those who have reached  $t$

$$\frac{\text{Prob}[\text{transition in next } \Delta t]}{\Delta t}$$

day or 8% in a year; for the 1-year interval centered on  $t = 59.25$ , the ID increases by about 0.024% per day – so little that even the fastidious Edmonds (see below) should not have been concerned – or just over 9% over the year. These almost-constant year-over-year hazard ratios of 1.08 or 1.09 for much of the age-range are similar to those that Gompertz observed in the material he studied, and bear out the log-linearity of mortality rates with respect to age that he termed a Law of Mortality.



Indeed, Clayton and Hills (1993, chapter 5 (Rates), p40) give it a yet-another name:

As the bands get shorter, the conditional probability that a subject fails during anyone band gets smaller. When a band shrinks towards a single moment of time, the conditional probability of failure during the band shrinks towards zero, but *the conditional probability of failure per unit time* converges to a quantity called *the probability rate*. This quantity is sometimes called the instantaneous probability rate to emphasize the fact that it refers to a moment in time. Other names are hazard rate and force of mortality.

Epidemiologists tend to first view it as an *empirical* quantity and define it using data. Indeed, Clayton and Hills *estimate* the rate parameter using the familiar incidence density measure:

In general, then, as the bands shrink to zero, the most likely (i.e., Maximum Likelihood) value of the rate parameter is

$$\frac{\text{Total number of failures}}{\text{Total observation time}}.$$

[...] This mathematical device of dividing the time scale into shorter and shorter bands is used frequently in this book, and we have found it useful to introduce the term *clicks* to describe these

very short time bands. Time can be measured in any convenient units, so that a rate of 1.11 per year is the same as a rate of 11.1 per 10 years, and so on. The total observation time added over subjects is known in epidemiology as the person-time of observation and is most commonly expressed as person-years. Because of the way they are calculated, estimates of rates are often given the units *per person-year* or *per 1000 person-years*.

One way to reconcile the two is to recognize that  $\text{Prob}[\text{transition in next } \Delta t]$  is the expected number of transitions<sup>7</sup> as a fraction of the number of candidates. Thus, just as with the Wikipedia definition, when we divide this probability by  $\Delta t$  to get what Clayton and Hill call the probability of failure per unit time, it becomes

$$\frac{\text{No. transitions in next } \Delta t}{\text{Ave. no. candidates}} \div \Delta t = \frac{\text{Ave. no. transitions in next } \Delta t}{(\text{Ave. no. candidates}) \times \Delta t},$$

which has the the same form as Clayton and Hills' estimator.

Although statisticians and epidemiologists understand that “time can be measured in any convenient units, so that a rate of 1.11 per year is the same as a rate of 11.1 per 10 years, and so on,” the next section shows that they sometimes forget how critical this point is, especially when one wishes to convert an (incidence-type) rate function into a risk.

---

<sup>7</sup>Since not all events in epidemiology involve movement from a more desirable (initial) state to a less desirable one, we use the more general term ‘transition’ instead of the term ‘failure.’

## 2 The units in which incidence rate, incidence density, hazard rate, and force of mortality are measured

Figure 2 depicts the water demand time curves for (and presumably, the degree of exclusively-television-viewing by the residents of) the city of Edmonton the afternoon (and the afternoon before) the U.S.A. ice-hockey team played the Canadian team in the gold-medal game at the 2010 Vancouver Winter Olympic Games. The graph has been viewed by more people than has Minard's classic portrayal of the losses suffered by Napoleon's army in the Russian campaign of 1812.<sup>8</sup>

While the behavior pattern is striking, one omission from the label "Consumer Water Demand, ML" for the vertical axis is notable. This author has shown this graph to a number of colleagues and outsiders, explaining if necessary that 'ML' is a measure of volume (it is short for 'megalitre' or millions of litres) aggregated over all consumers, and then asking if there is anything remarkable about the vertical axis. People with engineering-type training, or physicians who measure lung function, have very quickly responded that "it is missing a time-dimension." Curiously, epidemiology- and biostatistics-types whom he has posed this question have been slower to notice this omission. But even more interesting has been the split as to what they think the units

---

<sup>8</sup><http://www.edwardtufte.com/tufte/posters>

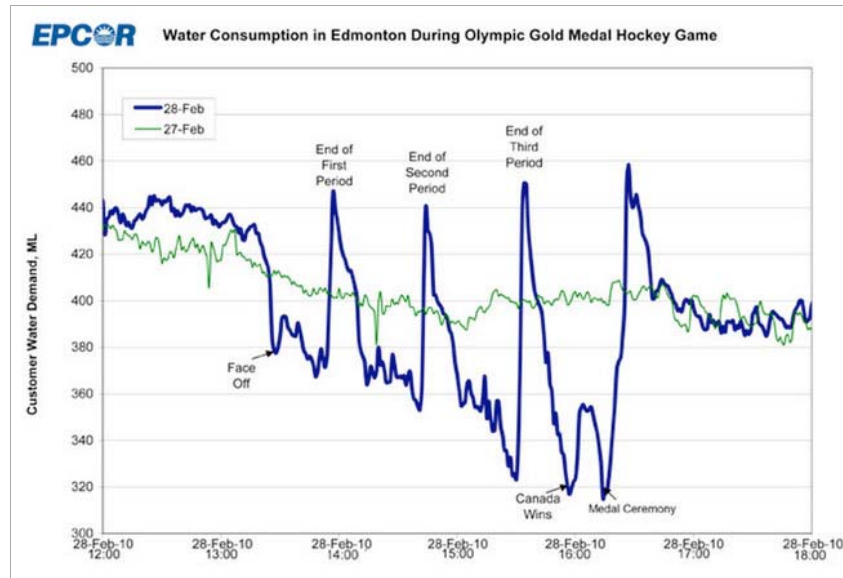


Figure 2: Minute-by-minute numbers of television-viewers of a major sports event. Q: What time units are missing from the label for the vertical axis?

of the missing dimension are. A number are quite confident that it must be ‘ML per-hour’ because “the time scale for the horizontal axis is marked off in hours”. Others are equally adamant that it must be ‘ML per-minute’ because “the graph fluctuates by the minute.” I leave it to readers to form their own opinions as to what the missing time unit is. Those who like to calculate can use the following data: 400ML is approximately 106 million U.S. gallons; the population of Edmonton is approximately 700,000 people; “the average Edmonton resident uses 230 litres/person/day for indoor and outdoor use.”<sup>9</sup>

To decide which unit most closely matches the reported usage, they will

<sup>9</sup><http://www.epcor.ca/>

probably compute the total volume over the 6 hours, by taking the (approximate) integral of (area under) the water-demand curve over this time-span. But in order to do so, each  $\Delta t$  on the t-axis must be in the *same* units as the ML/timeunit on the demand-axis: the volume for each subinterval is  $\frac{ML}{timeunit} \times \Delta t \text{ timeunits}$ . Thus, the total volume of demand for a given 15m interval is

$$Vol_{15min} = ? \frac{ML}{hour} \times 0.25 \text{ hours} = ? \frac{ML}{min} \times 15 \text{ min} = ? \frac{ML}{day} \times 0.01041\dot{6} \text{ days}$$

with  $?\frac{ML}{timeunit}$  denoting the average demand per time-unit over the interval in question.

The reaction that the demand must be ML per minute because the time scale is in minutes is similar to the one which says that if we are to graph the velocity of a car over a period of minutes, we have to measure the speed in miles per minute rather than in mph– or that we cannot express heart rate in beats per minute if we only measure for 15 seconds. We can scale the velocity to any time unit we wish, but if the integral is to represent the total distance travelled, we need to calculate the distance travelled in each different subinterval of time as the (average) distance per time unit over that subinterval  $\times$  the time-length of that subinterval – with the time-duration expressed in the same units as was the velocity.<sup>10</sup> This issue of time units

---

<sup>10</sup>For a striking example of improper use of units, and by the confusion caused by the statement that the “venous thromboembolic incidence was 3.6% and 1.5% in the first and second weeks postpartum, respectively, similar to the 2% to 5% incidence of symptomatic venous thromboembolism after elective hip replacement in patients not receiving prophylaxis”

becomes paramount in section 4, when we convert an incidence function into a risk.

### 3 Edmonds, the continuous force of mortality, and the concept of a person-moment and a person-year

Benjamin Gompertz used the word ‘*intensity*’ of mortality in his 1825 article.<sup>11</sup> We believe that the first person to use the term ‘*force*’ of mortality in writing was T.R. Edmonds, a political economist and actuary, and a neighbor and collaborator of William Farr (Eyler 2002; Turner and Hanley, 2010). Edmonds put the term in italics and in quotes in the first paragraph of his 1832 book. He begins his theoretical treatment with the words (emphasis ours)

The force of mortality at any age is measured by the number of

---

laxis” (when “only 105 maternal cases of venous thromboembolism were diagnosed during pregnancy or postpartum in 50 000 births”), see the correspondence regarding “Incidence of Pregnancy-Associated Venous Thromboembolism and family history as major risk factors” begun by MacCallum et al. in the Annals of Internal Medicine, 21 March 2006 Annals of Internal Medicine Volume 144(6). pp 453-460.

For an striking example of how different units can make an incidence function look larger or smaller, epidemiology students might wish to convert the army losses in Napoleon’s Russian campaign into incidence densities, as a function of elapsed time ( $d^{-1}$ , or  $m^{-1}$ , or  $y^{-1}$ ) or distance ( $\text{mile}^{-1}$ , or  $\text{Km}^{-1}$ , or  $^{\circ}\text{long}^{-1}$ ). The data are available at <http://www.math.yorku.ca/SCS/Gallery/re-minard.html>.

<sup>11</sup>Linder (1936) tells us that in 1765, “mathematician Johann Heinrich Lambert, (1728-1777) was the first to direct attention to what he calls ‘Lebenskraft,’ that is, the force of vitality or the reciprocal of the force of mortality”.

deaths in a given time, *out of a given number constantly living.*

The given time has been here assumed to be one year, and the given number living to be one person;

Whereas he defined the force of mortality as “the quantity of death in one year for a unit of life at the assumed age” he conceded that “the force is changing continually” and so he gives a more hypothetical definition “the quantity of death on a unit of life which *would* occur by the action of this force *continued uniform* for the space of one year. Edmonds employed infinitesimal calculus to use the “relation of Dying to Living for *large* intervals of age to deduce and interpolate the relation corresponding to *small* intervals of age”:

“[S]ince this relation for annual intervals is continually varying, it is manifest, that the same principles which have led to the conclusion, that the variation is continued and annual, must lead to the conclusion, that the variation is monthly, and also to the conclusion, that the variation is diurnal, and even momental.<sup>12</sup>

It may be assumed, therefore, that all Tables of Mortality represent the relation of Dying to Living as changing continuously, - that this relation is never the same for any two successive instants of age. I have used the term ‘force of mortality,’ to denote this relation at any *definite moment* of age. It would evidently

---

<sup>12</sup>The word ‘person-moment’ is used in Miettinen OS. Etiologic research Needed revisions of concepts and principles Scand J Work Environ Health 1999;25 (6, special issue):484-490.

be improper to use this term to express the relation of Dying to Living in yearly intervals of age; for the force of mortality at the beginning, at the middle, and at the end of any year of age, are all different.”

Part II will exploit Edmond’s idea of “one person ‘constantly living’ for one year” (what today would be called a dynamic population with a constant size of 1) to more simply and heuristically derive the fundamental relationship between an incidence function and the cumulative incidence proportion (risk) that seems to have been neglected or made unnecessarily complicated in modern textbooks.



## References

- Ayas NT, Barger LK, Cade BE et al. Extended Work Duration and the Risk of Self-reported Percutaneous Injuries in Interns. *JAMA*. 2006; 296: 1055-1062.
- Schröder FH, Hugosson J, Roobol MJ, et al. ERSPC Investigators. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009 Mar 26;360(13):1320-8. Epub 2009 Mar 18.
- Ridker, P. et al. *nejm* nov 20, 2008 Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated CRP.
- Vandenbroucke JP, Pearce, N. Incidence rates in dynamic populations. *International Journal of Epidemiology*. To appear.
- Miettinen, O. S. (1976) Estimability and estimation in case-referent studies. *Am. J. Epidem.*, 103, 226-235.
- Gompertz, B. (1825) On the nature of the function expressive of the law of human mortality, and on a new mode of determining life contingencies. *Phil. Trans. R. Soc. Lond.*, 115, 513-583.
- Edmonds, T. R. (1832) *The Discovery of a Numerical Law regulating the Existence of Every Human Being illustrated by a New Theory of the Causes producing Health and Longevity*. London: Duncan. Available as an on-line digital version at <http://books.google.com>.
- Clayton D, Hills M. *Statistical Models in Epidemiology*. Oxford University Press, 1993.
- Eyler JM. Constructing vital statistics: Thomas Rowe Edmonds and William Farr, 1835-1845. *Soz.- Präventivmed*. 47 (2002) 006-013, 2002

Turner EL and Hanley JA. Cultural imagery and statistical models of the force of mortality: Addison, Gompertz and Pearson. *J. R. Statist. Soc. A* (2010) 173, Part 3, pp. 483-499.

Vandenbroucke JP, Pearce, N. Incidence rates in dynamic populations. *International Journal of Epidemiology*. To appear.