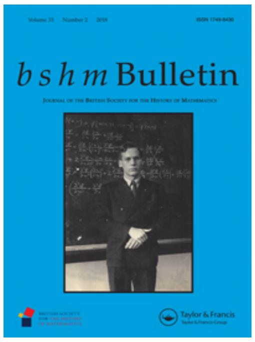

Taylor & Francis
Taylor & Francis Group

BSHM Bulletin: Journal of the British Society for the History of Mathematics

ISSN: 1749-8430 (Print) 1749-8341 (Online) Journal homepage: https://www.tandfonline.com/loi/tbsh20

Edward L. Kaplan and the Kaplan-Meier Survival Curve

Lukas J A Stalpers & Edward L Kaplan

To cite this article: Lukas J A Stalpers & Edward L Kaplan (2018) Edward L. Kaplan and the Kaplan-Meier Survival Curve, BSHM Bulletin: Journal of the British Society for the History of Mathematics, 33:2, 109-135, DOI: 10.1080/17498430.2018.1450055

To link to this article: https://doi.org/10.1080/17498430.2018.1450055

Published online: 18 May 2018.

BSHM Bulletin, 2018
Vol. 33, No. 2, 109–135, http://dx.doi.org/10.1080/17498430.2018.1450055

Edward L. Kaplan and the Kaplan–Meier Survival Curve

Lukas J A Stalpers
University of Amsterdam, The Netherlands

Edward L Kaplan
University of Minnesota Medical School, USA

In June 1958, Edward L Kaplan (1920–2006) and Paul Meier (1924–2011) published an innovative statistical method to estimate survival curves when including incomplete observations. The Kaplan–Meier (KM) method became the standard way of reporting patient survival in medical research. For example, the KM method is used in more than 70% of clinical oncology papers. With 44,319 Web of Science® citations as of November 2017, the report has become the most-cited statistics publication in the scientific literature. Part I of this report describes the KM method, its strengths and limitations, and the history and evolution of the method. In Part II we recount the biography of the remarkable mathematician Edward L Kaplan, PhD, and his unique contributions during the formulation of the KM method, as well as his contributions to science during his unique and productive career.

Introduction

In the June 1958 issue of the *Journal of the American Statistical Association* (*JASA*) Edward L Kaplan and Paul Meier published a manuscript describing an innovative statistical method to estimate survival rates when there are incomplete survival observations included in the data base (Kaplan and Meier 1958). With 44,319 Web of Science® citations as of November 2017, the paper is the eleventh most cited scientific paper of the modern era: the second most cited paper in mathematics, ranking just below a bio-informatics paper by Thompson *et al.* (1994), and the most cited statistical paper (Garfield 1983; Ryan and Woodall 2005; van Noorden *et al.* 2014). Moreover, the Kaplan–Meier method has become the standard way to report patient survival in medical and epidemiological studies, particularly in cancer research (Cox 1972; Peto 1976; van Noorden *et al.* 2015).

Edward Lynn Kaplan (1920–2006) and Paul Meier (1924–2011) were graduate students in mathematics at Princeton University, and had the same PhD mentor, Professor John Tukey (1915–2000). Paul Meier's contribution to this important work was frequently acknowledged during his lifetime; he occupied key academic positions in medical statistics, and was a board member of many important statistical and public health organizations. Paul Meier received several distinguished awards for his contribution (Marks 2004). At the time of his death, his life was commemorated in obituaries in major national newspapers and scientific journals.

What about the first author? Edward Lynn Kaplan, by contrast, has largely been overlooked in academic accounts of modern mathematics and statistics and the profound influence of his contribution—including providing a basis for modern medical advances—has seldom been recognized. We asked: why?

Survival analysis: life, death and twilight states

'Patient survival is generally accepted as the principal criterion for measuring the effectiveness of treatment in cancer' (Ederer 1961) ... and also in other life-threatening diseases (Willett and Singer 1991; Cole and Hudgens 2010; Stel *et al.* 2011; May and McKnight 2017). Cancer treatment results are typically reported as survival rates after 1 year, 5 years or 10 years. For the last four decades, the survival rate $S(t)$ at time t has been predominately calculated by the Kaplan–Meier product–limit method, and presented as a Kaplan–Meier 'staircase' survival curve (see example, Figure 1 in next section). Before the KM method was utilized, the 5-year chance of survival of a group of patients could only be reliably calculated by following every patient in the cohort for at least the full 5 years, using life-table methods (Hayward 1899a,b, 1900, 1902a,b; Greenwood 1926; Berkson and Gage 1950).

However, in a typical survival study, the accrual of a sufficient number of patients with a specific illness usually takes many years. For example, for a study period of 10 years, many patients can only be followed for less than 5 years, and others may be lost to follow-up because they cannot be located. These patients, alive or lost to follow-up, are technically called 'censored' or 'incomplete' observations. These censored observations must be included in estimating the survival rate $S(t)$. This problem with incomplete observations in survival studies had already been recognized in the nineteenth century (Hayward 1899a,b, 1900; Greenwood 1926). Greenwood (1926) already remarked that if living patients with a shorter observation time than the duration of the study were removed from the cohort (that is, 'censored' in the grammatical and legal sense), then the survival

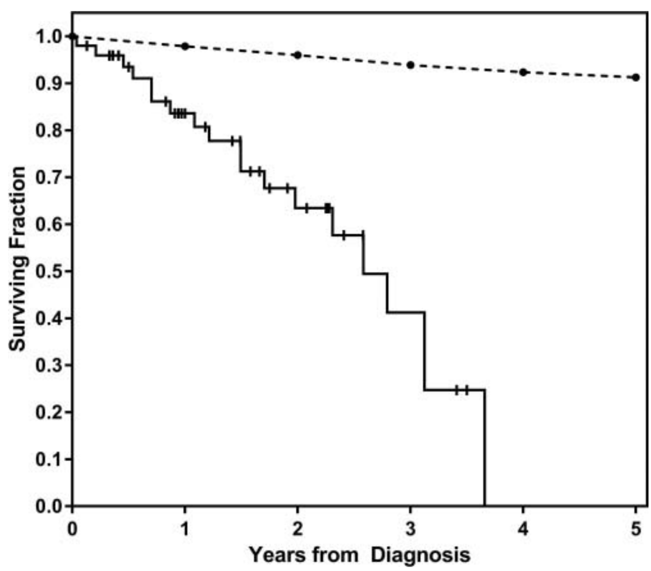

Figure 1. Re-analysis of the survival of the first fifty breast cancer patients operated upon by William S Halsted between 1889 and 1894 using the Kaplan–Meier method (solid line; source data in Halsted 1894). For comparison, we added the relative survival curve of US breast cancer patients in 2009 (dotted line; source: SEER 1975–2014 data by Howlader *et al.* 2017)

estimate would be too low (Greenwood 1926). On the other hand, if these living patients with an observation time shorter than the duration of the study were assumed to remain alive until the end of the study, the survival estimate would be too high.

Kaplan and Meier offered a solution for this significant problem. In the early 1950s, when both were graduate students in the Department of Mathematics at Princeton University, each separately worked on the problem of survival analysis under the supervision of their graduate advisor, Professor John Tukey. Heeding the advice of the then Editor of the *JASA*, the original two separate manuscripts were merged and published as 'Nonparametric estimation of survival from incomplete observations' (Kaplan and Meier 1958). In that seminal paper, they addressed the problem of censored patients by dividing the survival curve into discrete time intervals defined by the time of every subsequent patient death ('event'). Next, the mortality risk within each interval was calculated as the proportion of the number(s) of deaths in that interval divided by the number of living patients at the beginning of the interval. Finally, a chain of survival rates was calculated by linking the subsequent mortality risks by a 'product–limit estimator'. (See S1: Supplementary Mathematical Details.)

To mathematicians, the Kaplan–Meier method is relatively simple. Furthermore, today every physician or basic scientist can calculate Kaplan–Meier survival curves using one of many statistical software packages (for example, SAS, SPSS, Stata, GraphPad Prism or R). Most of the work for the clinical researcher lies in collecting three pieces of information for each patient: (1) The first date of observation (in patient studies typically the date of disease diagnosis or treatment); (2) the last date of observation (typically the date of death or the last date that a patient was seen alive); and (3) the vital status on that last date, whether that is an 'event' (typically death) or a 'censored observation' (in patient studies typically 'alive' or 'lost to follow-up').

Despite the mathematical simplicity of the product–limit estimator, it took Kaplan and Meier several years to fine-tune their manuscript. In a later interview with Eugene Garfield for *Current Contents* in 1983, Kaplan recounted the history of the manuscript (Garfield 1983):

> This paper began in 1952 when Paul Meier at Johns Hopkins University ... encountered Greenwood's paper on the duration of cancer. A year later at Bell Telephone Laboratories, I became interested in the lifetimes of vacuum tubes in the repeaters in telephone cables buried in the ocean. When I showed my manuscript to John W Tukey, he informed me of Meier's work, which already was circulating among some of our colleagues. Both manuscripts were submitted to the Journal of the American Statistical Association, which recommended a joint paper. Much correspondence over four years was required to reconcile our differing approaches, and we were concerned that meanwhile someone else might publish the idea.

The latter was not an unjustified concern. Joseph Berkson and Robert P Gage from the Mayo Clinic in Rochester, Minnesota were working on the same topic. They were very much aware of the shortcomings of their attempts to give actuarial survival estimates, and were thinking of a solution (Berkson and Gage 1950, 1952; Berkson 1954). Moreover, later (August of the same year that Kaplan and Meier published their paper in *JASA*), Sidney Cutler and Fred Ederer from the National Cancer Institute in Bethesda, MD, published an alternative method for calculating *actuarial* or *life-table survival curves* (Cutler and Ederer 1958; Garfield 1979; Green 1997).

According to Paul Meier, in an interview with Harry Marks in 2004, both he and Kaplan were working independently on the same topic, and Kaplan was first to submit his manuscript to *JASA*. Meier stated that Kaplan had attributed some part of the combined work 'on the bias of estimates' to Meier (Marks 2004). Still in 2004, Meier had difficulty in fully acknowledging Kaplan's contribution to the collaborative effort which resulted in the final merged manuscript:

> Well, I got pretty mad and I called John (Tukey). Ultimately, I wrote to the editor of JASA about the situation; by that time, I had estimated the means and stuff like that and not just Kaplan–Meier alone, and I offered JASA this as a second paper. The editor wrote back that 'our readers wouldn't like it if you separate this into two papers. I'd rather you both got together and wrote one paper.' I swallowed hard, and I guess Kaplan swallowed hard as well. So, we worked quite hard and at one place he solved a problem that I couldn't solve; other cases I solved problems he couldn't.
>
> (Marks 2004)

Despite a concerted effort, we have not been successful in locating the two original separate manuscripts or in more specifically defining the contribution of each manuscript to the final merged and published version of 1958.

The Kaplan–Meier test applied to William Halsted's mastectomy patients

An example of how important the resulting KM method is can be shown by a reanalysis of the original survival data in the seminal paper of the renowned surgeon, William S Halsted (1852–1922), about radical mastectomy for women with breast cancer (Halsted 1894). The so-called Halsted procedure was the standard for breast cancer surgery for almost 80 years, and many thousands of women around the world had this very extensive surgical procedure in an attempt to save their lives. Would the procedure have been so widely used if Halsted could have presented his data using a Kaplan–Meier curve?

In 1894, Halsted published the original fifty case-histories of women with breast cancer operated on by him between 14 June 1889 and 2 February 1894, followed until early March 1894 (Halsted 1894). Halsted reported—for those days—a spectacularly low local recurrence rate of '6 Per Cent' (three cases), and only eight cases (16%) with a 'Regionary Recurrence'. According to that original report, only nineteen patients had died, which gives a 62% crude survival rate. Since a surgeon with Halsted's stature did not require peer review, the manuscript was immediately accepted and published. Critical remarks about biased results due to the short and incomplete follow-up appeared only much later (Korteweg 1903; Keynes 1929).

Figure 1 shows the survival curve of these same patients estimated by the Kaplan–Meier method, which corrects for incomplete observations due to short follow-up. For comparison we added the most recent survival curve for breast cancer patients from the US SEER-database (Howlader *et al.* 2017). Here it becomes clear that only 41% of Halsted's patients survived more than 3 years, and that few, if any, would have survived more than 5 years. The main contribution by Halsted was not that he cured more patients than others, but that more patients left the hospital alive thanks to fewer post-surgical deaths from infections: the nurses and doctors in Halsted's clinic were first to use rubber gloves in the operation room, starting with his nurse who had developed contact dermatitis from the disinfectant chemicals (Lathan 2010). It was serendipity that rubber gloves also prevented the usually fatal wound infections.

The acceptance of the Kaplan–Meier method by the scientific community

In a 1990 overview by Garfield of the hundred most-cited science papers between 1945 and 1988, the Kaplan–Meier paper was at rank number 55: already the most cited statistical paper. A 2005 analysis of 'The most cited statistical papers' by Ryan and Woodall (2005) showed that the number of citations of the 1958 paper since 1990 indicated that its ranking had further increased (Ryan and Woodall 2005, 2011; see S3). In 2014, the Kaplan–Meier paper's citations further increased (38,600 citations), bringing it to eleventh position among the most cited papers in the Web of Science® 'Top 100 papers' (van Noorden *et al.* 2014). Figure 2 presents a Web of Science® analysis of the citation curve for the Kaplan–Meier paper, which reached a peak of 1790 citations during the year 1997.

The citation history in Figure 2 shows another remarkable feature: although the Kaplan–Meier paper was published in 1958, it went largely unnoticed until 1969. We reviewed the first one hundred papers that cited the Kaplan–Meier paper. This revealed that until 1969, the paper was cited on average only once a year.

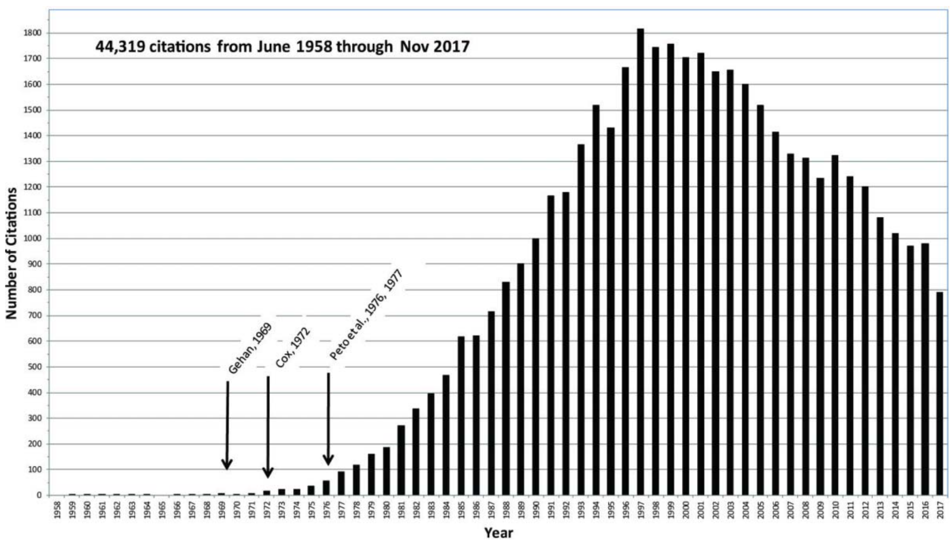

Figure 2. Web of Science® citation history of the Kaplan–Meier publication since 1958 as of November 2017, totalling 44,319 citations

The first three of those one hundred papers are worth mentioning in more detail, since all three authors later became leading figures in biostatistics. Peter Armitage (born 1924) was the first who cited Kaplan and Meier in a paper on 'The Comparison of Survival Curves' (Armitage 1959), followed in the same year by a paper by (later Sir) David R Cox (born 1924) describing the first step towards what would be published fourteen years later as the proportional hazards model, for which the Kaplan–Meier curve became the underlying survival model (Cox 1959, 1972). Next, in 1961 Fred Ederer (born 1926) cited the Kaplan and Meier 1958 publication in a paper in which he proposed a simplified version of the Greenwood standard error for survival curves (Ederer 1961a). As previously mentioned, Cutler and Ederer in 1958 published a different actuarial or life-table method to calculate survival considering incomplete data (Cutler and Ederer 1958). Of lasting importance is Ederer's contribution to the calculation of relative survival; particularly the method by Ederer and Heise from 1958 (also known as Ederer II) (Ederer and Heise 1959; Cho *et al.* 2011). Ederer II is used by both the main American and European cancer registries to report internationally comparable 5-year relative survival rates, although the available software packages use the Kaplan–Meier curve as the underlying model, and not the life-table method by Cutler and Ederer (Ederer and Heise 1959; Ederer *et al.* 1961b; Dickman and Coviello 2015; Pohar and Stare 2006; Cho *et al.* 2011).

As Julian Peto (1984) remarked, the life-table method by Cutler and Ederer (1958) was actually slightly more accurate than the Kaplan–Meier method (1958) for estimating survival with incomplete observations compared to survival with complete observations (Peto 1984). In due course, many more statisticians proposed improvements to both the KM product–limit estimator and the Greenwood confidence interval for the KM survival curve (see S2: Limitations of the Kaplan–Meier method).

The breakthrough in citations of the Kaplan–Meier paper started in 1969 with a publication by Edmund A Gehan, the chief of biostatistics at MD Anderson Cancer Center in Houston, TX (Gehan 1969). Gehan is particularly known for his contribution of the general Wilcoxon test for comparing survival curves (Gehan 1965). His 1969 paper was an important methodological contribution on survival statistics, in which Gehan introduced the KM product–limit estimator arguing that the then current 'life table method is nearly equivalent to this in large samples' (Gehan 1969). Small sample size is particularly a problem in clinical studies, and therefore in 1971, the MD Anderson group applied the Kaplan–Meier method to the then available cancer survival data in at least four papers (Bodey *et al.* 1971; Coltman *et al.* 1971; Fuller *et al.* 1971; Luce *et al.* 1971). In the first of those four clinical papers, a study by Coltman *et al.* (1971) on the results of chemotherapy for solid tumors, we also found the initial graphic representation of the Kaplan–Meier survival estimates by the typical staircase curves (Figure 3).

Although Kaplan had originally developed his part of the KM method explicitly for non-biological purposes (for example, the survival of vacuum tubes at Bell Laboratories) the noteworthy 1969 contribution of Gehan, using the KM method to describe the survival of patients with cancer, represented its introduction into the medical literature (see Figure 3) and was the beginning of its exponential rise in literature citations.

A second major impetus for the increasing utilization of the Kaplan–Meier method was the seminal paper of Sir David R Cox (1972) on 'Regression models and

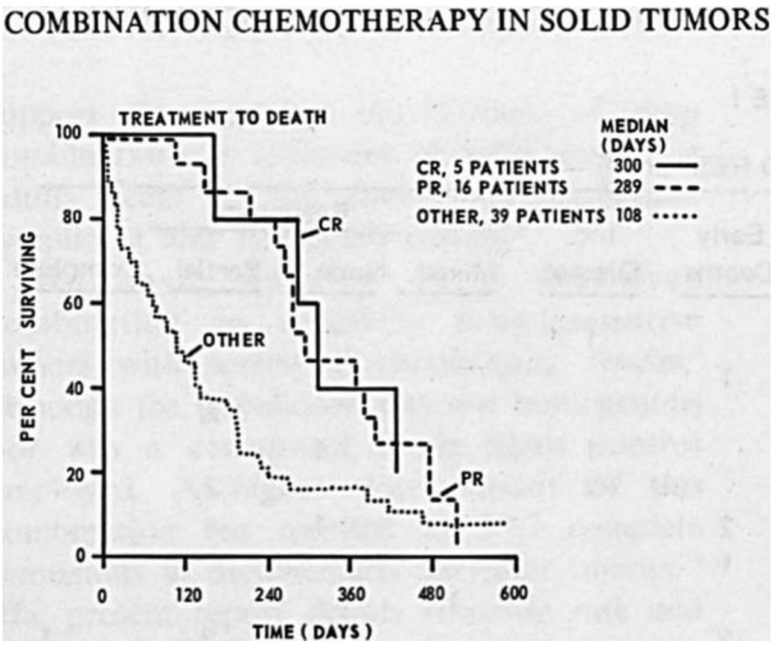

Figure 3. Survival after chemotherapy for solid tumours stratified by tumour response. This graph taken from the paper by Coltman *et al.* (1971) from MD Anderson Cancer Center, with Edward Gehan as the biostatistician, is probably the first typical 'staircase' medical survival representation of Kaplan–Meier survival curves. (Reproduced with permission from Elsevier®)

life tables'. After the Kaplan–Meier paper, the Cox paper ranks as the second most cited statistics paper (van Noorden *et al.* 2014). Cox was clear about the importance of the Kaplan–Meier paper (Cox 1972):

> Life tables are one of the oldest statistical techniques and are extensively used by medical statisticians and by actuaries. Yet relatively little has been written about their more formal statistical theory. Kaplan and Meier (1958) gave a comprehensive review of earlier work and many new results. ... The present paper is largely concerned with the extension of the results of Kaplan and Meier to the comparison of life tables and more generally to the incorporation of regression-like arguments into life-table analysis.

Both the Kaplan–Meier method and the Cox method were further boosted by two tutorial papers by the distinguished physician Sir Richard Peto *et al.* (1976, 1977) on the 'Design and analysis of randomized clinical trials requiring prolonged observation of each patient'.

In hindsight it is remarkable that Cox, a prominent British statistician from Imperial College London, who had close connections with statisticians Cutler, Mantel, Haenszel, and Ederer from the National Cancer Institute, did not select the smooth 'monotonic' life-table method by Cutler and Ederer (1958) from the National Cancer Institute, but instead chose the stepwise Kaplan–Meier curve. The explanation probably can be found on page 190 of Cox's paper:

> Alternatively we may restrict $\lambda_0(t)$ qualitatively, for example by assuming it to be monotonic or to be a step function (a suggestion of Professor J W Tukey)....
>
> (Cox 1972)

Cox accepted Tukey's suggestion and adopted the Kaplan–Meier step function, and probably did so as early as 1959 (Cox 1959). As mentioned earlier, John Tukey was at Princeton University and was the PhD mentor of both Kaplan and Meier. He also had a collegial relationship with Cox:

> John Tukey was the first person I met the first time I visited the US. He met my family and me at the dock side in 1955 when we arrived by ship as one did in those days. I had possibly hundreds of conversations with him in the subsequent years. ...(O)ne of his many strengths was the ability to comment searchingly on many things.
>
> (Cox, personal communication, August 2016)

Moreover, Gehan visited Cox during the period when the latter was working on the proportional hazards method, providing the opportunity to discuss the principles contained in the KM method:

> Gehan I knew when he was a doctoral student in North Carolina and he spent a year in London with me in about 1970.
>
> (Cox, personal communication, August 2016)

Has there been a true decline in the use of the Kaplan–Meier method?

The number of KM citations peaked in 1997. Figure 2 shows a gradual decline in the number of yearly citations to 935 in 2016. This is still a very respectable number, that has resulted in an increase in the Web of Science® citation ranking of the paper from the 55th position in 1983 to the eleventh position in 2014, as well as becoming the most cited statistical paper in 2016. But does the decreasing citation number indicate a declining impact of the KM method?

We reviewed all 47 'Clinical Investigations' (complete papers) published during the first three issues of 2017 (January, February and March 2017) of *The International Journal of Radiation Oncology, Biology and Physics* (*IJROBP*), a leading journal in radiation oncology. Survival was an explicit end-point in 34 papers (72%). The KM method was used to estimate survival in all of them, but it was only mentioned in the Methods section in 23 (67%) papers. The paper from 1958 was cited in only 5 out of 34 papers (15%). Four other papers cited the Cox model or an adaptation of it (Cox 1972; Lin and Wei 1989; Fine and Gray 1999). If the *IJROBP*, a journal with high standards regarding the correct use of methodology, is taken as representative for all medical journals, the Web of Science® citation rate of the KM paper for 2017 may be as much as a seven-fold underestimation of the actual frequency with which the method is used in published medical studies.

We do not wish to advocate that the KM method should be cited in every paper in which the method is used. We simply wish to demonstrate that the current citation rate of the KM paper is likely an incomplete representation of the continuing popularity of the method among medical scientists and clinicians. Indeed, this evidence confirms that the Kaplan–Meier method has become a 'household' statistical method.

Edward L Kaplan's life and achievements

Above, we described the development of the KM method. Next we review the biography of Edward L Kaplan and his essential contributions to the KM method. We describe the previously unrecognized life and fate of this exceptionally prodigious and sensitive mathematician, who drastically broke with research and has therefore been overlooked by scientific history. He was not only the co-discoverer of the Kaplan–Meier method, but our recent research has revealed his significant—and acknowledged by colleagues—contributions while at the US Navy Ordnance Laboratory during World War II, and later at the Bell Laboratories in New Jersey and the Lawrence Livermore Laboratories in Berkeley, California. A majority of his almost sixty scientific manuscripts relating to the war effort have not been previously recognized because they were classified during and immediately after the war and therefore could not be submitted for publication. After developing personal concerns regarding nuclear research in 1961, and then a divorce, Kaplan became a less productive contributor to academic mathematics. His important contributions to mathematics deserve wider public and academic acknowledgment.

Academic biography[1]

Edward Lynn Kaplan (born Philadelphia, Pennsylvania, 11 May 1920; died Corvallis, Oregon, 26 September 2006) graduated from Swissvale High School in Swissvale, Pennsylvania in 1937. His parents were Eugene V Kaplan (1887–1977) and Frances Rhodes Kaplan (1891–1978). Edward Kaplan attended the Carnegie Institute of Technology (CIT) from 1937 to 1941 and graduated with a Bachelor's degree in mathematics in 1941. He received 'Second honors in a class of 200' (citation from US National Archives 2016). During his college years, he was president of the mathematics club and sang in the men's glee club. An extraordinary three times—in 1939, 1940 and 1941—he was one of the five honorees in the prestigious nationwide William Lowell Putnam Mathematical Competition (American Mathematical Association 2016a) (Figure 4). Through being elected a Putnam Fellow, he also was offered a Westinghouse scholarship to Harvard University which he declined:

> I was awarded the Putnam Prize Scholarship in mathematics at Harvard; the war prevented my accepting.
>
> (Kaplan 1999; US National Archives 2016)

Edward Kaplan was elected as a member of three scholastic societies: Phi Kappa Phi (recognition and encouragement of superior scholarship), Sigma Xi (scientific researchers), and Tau Beta Phi (engineering students).

From June 1941 to August 1948, Kaplan worked at the United States Naval Ordinance Laboratory, Whiteoak, Maryland. After the war, in 1948, he went to

[1]See also: S6: Chronology of the Life of Edward Lynn Kaplan.

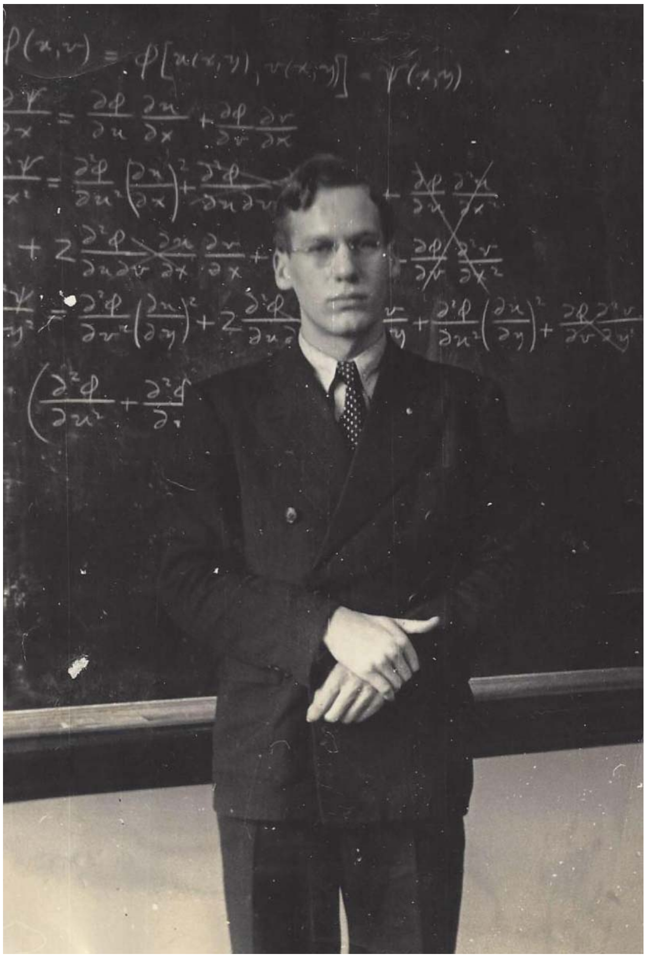

Figure 4. Edward L Kaplan as a sophomore at Carnegie Institute of Technology at age 19 years. The photograph was taken on the occasion of the presentation of the Putnam prize to Kaplan in 1939 (see text) (Source with permission: Kaplan family archive)

Princeton University as a PhD student in the department of mathematics. This was the same year and the same department as John Nash, Jr (Nobel laureate in 1994; also a Carnegie Institute of Technology graduate). Their mathematics tutor at Carnegie, Professor Joseph B Rosenbach, wrote in *The Carnegie Tartan* (the university student newspaper) of 20 April 1948:

> John Nash and Edward L Kaplan, a graduate from the department of mathematics in 1941, are the two best students I have ever had in the past 30 years of college teaching. They have the quickness of understanding, the originality, and the capacity for seeing the inter-meaning of argument that are unrivalled in my experience.

Following graduation from Carnegie, Kaplan worked at the Naval Ordnance Laboratory where his department chief was Dr John V Atanosoff (1903–95), renowned as the inventor of the first electronic computer. In the margins of a questionnaire from 1948 in support of Kaplan's application for a promotion at the Naval Ordnance Laboratory, Atanosoff attested:

> This man is the best B.S. level mathematician I have ever seen. His ability far excels that of many men with the PhD degree.
>
> (US National Archives 2016)

In the summer of 1949, Kaplan briefly worked for the United States Air Force. He finished his PhD thesis in November 1950, on 'Infinite permutations of stationary random sequences' (Kaplan 1950b, 1955a). His mentors at Princeton were Professors John W Tukey (1915–2000), and Samuel S Wilks (1906–64) (Mosteller 1964; Schultz and Eisenberg 2000).

After graduation from Princeton, Kaplan began working as a member of the staff of Bell Telephone Laboratories in Murray Hill, NJ, where he remained until 1957. From there he moved to the Computation Division of the University of California at Berkeley, Lawrence Radiation Laboratory, Livermore, CA, directed by Dr Sidney Fernbach. At Livermore he worked on Monte Carlo simulations required for the development of the hydrogen bomb. However, on 25 July 1961 he applied to the Oregon State University, and moved to Corvallis, Oregon to become a professor of mathematics there. Edward Kaplan retired in 1980; he died in Corvallis, Oregon on 26 September 2006 at the age of 86 after a prolonged debilitating illness (Kaplan 1961a,b, 1999; Smythe 2006).

Accomplishments of Edward L Kaplan

As discussed in detail above, Kaplan is specifically recognized for his 1958 publication with Paul Meier (1924–2011) on what became known as the KM survival curve (Kaplan and Meier 1958). The Web of Science database lists only this paper as his published output.

However, in his 1961 application letter to Oregon State University, he also included the titles of six peer reviewed papers (Gilfillan and Kaplan 1941; Kaplan 1946a, 1948a, 1950a, 1952a,b) and his PhD thesis (Kaplan 1950b, 1955a), and mentioned 'several reports' that were unpublished because they were categorized as 'classified' by the United States Government.

During our research, it became clear that here Kaplan was modestly referring to another fifty classified (at that time) reports that he authored based on his work for the Department of the US Navy (Naval Ordnance Laboratories near Washington, DC), the Bell Labs, and the Livermore Laboratories at the University of California at Berkeley, which were the main sponsors of his research projects between 1941 and 1961 (Kaplan 1942a,b,c, 1943a–k, 1944a–c, 1945a–f, 1946b, 1947a–f, 1948a–g, 1953, 1955b,c, 1956, 1958a,b, 1961a,b; Kaplan and Mooney 1943; Kaplan and Slawsky 1947; Kaplan and Yost 1948; Kaplan and Terry 1954). Our supplementary file S5 presents a list of his sixty-one publications, reports and manuscripts published between 1941 and 1999 (United States National Archives 2016) .

In his first paper, in 1941, he was the co-author with Gilfillan on a manuscript discussing mechanical problems in induction motors (Gilfillan and Kaplan 1941).

The next three publications about elliptic integrals resulted from Kaplan's classified research during World War II for the United States Navy (Kaplan 1946a, 1948a, 1950a). Elliptic integrals are complex mathematical functions, which cannot be expressed in terms of elementary functions. Initially developed as an abstract mathematical exercise, elliptic integrals became practically relevant in the description and prediction of complex movements: such as determining the position of a moving ship on the sea using radar, and guiding a torpedo to it. Nowadays, these complex mathematical problems can be easily solved by computer software, but in the 1940s the Navy used tables such as those compiled by Kaplan.

The solutions of elliptic integrals remained an important element in Kaplan's later work, including his Princeton PhD thesis on 'Infinite permutations of stationary random sequences' (Kaplan 1950b, 1953). However, influenced by the work of statistician Joseph L Doob (1910–2004), he sought solutions for complex operational processes more in terms of random processes and probabilistic transitions (Doob 1948, 1953; Kaplan 1950b, 1953, 1955a). From references in his PhD thesis, it is clear that Kaplan was in close contact with Doob, and that Kaplan had access to the manuscript of Doob's *Stochastic processes*, that would be published in 1953 and became a classic in probability theory (Doob 1948, 1953). In fact, the seminal 1958 Kaplan–Meier paper can be seen as studying a simple stationary random sequence, reduced to three states: alive, dead, and 'lost to follow-up'. The first design for such a reduced three-state model can be found in an internal report at the Bell Labs (Kaplan 1953).

We identified this report on 'The reduction of incomplete life test data' as the early manuscript that Kaplan might have shown to Tukey in 1953 (Kaplan 1953; Garfield 1983; Marks 2004). The two striking differences compared with the final paper—presumably indicating contributions made by Meier—were the focus on the life-span of vacuum tubes rather than survival-time of cancer patients, and the absence of the Greenwood standard error for the survival rate.

Although the original manuscript by Paul Meier appears to be lost, there are several reasons to think that Meier's main contribution to the KM paper was the use of Greenwood's (1926) calculation of the standard error. First, Meier's PhD thesis was largely concerned with variance and standard errors of the mean (Meier 1951, 1953). Next, in a 1983 interview, Kaplan identified Greenwood's publication as Meier's inspiration for (his contribution to) the KM paper:

> This paper began in 1952 when Paul Meier at Johns Hopkins University ... encountered Greenwood's paper on the duration of cancer
>
> (Kaplan, quoted in Garfield 1983).

Finally: Meier himself confided during an interview in 2004 that he considered estimation of variance or what he then called 'bias' the most essential part of the KM paper:

> Berkson from the Mayo Clinic had written a paper about it [survival analysis], but he hadn't estimated the variance. Somebody asked me how to do it, and I said, 'Oh, that's very hard: you have to do this and that and ... '. Then one of my

> colleagues showed me a Major Greenwood paper on it that opened my eyes quite a bit ... .

and:

> I didn't know Kaplan, but I wrote to him about my ideas, including the notion that if you only have two instead of five years observations on some patients, the estimates may be quite biased. Kaplan had thought the estimates were unbiased. He wrote back and said that he had credited the idea of biased estimates to Meier, and sent his paper off to JASA
>
> (Meier, quoted in Marks 2004, 132, 133).

The Greenwood confidence interval has been replaced since by more accurate ones (as described in S2: 'Limitations of the Kaplan–Meier method') (Rothman 1978; Borkowf 2005; Miettinen 2008). But Kaplan's product–limit estimator still stands.

The problem of dealing with incomplete data kept Kaplan busy for the next few years at Bell Laboratories, and resulted in at least three unpublished internal reports (Kaplan 1954, 1955a, 1956). His work on Monte Carlo simulations for the hydrogen bomb at Livermore was written up in at least three internal reports and a conference paper, but these, too, were never published in scientific journals (Kaplan 1958a,b, 1961a,b). There is no evidence that he submitted any of these manuscripts for publication.

Moreover, during his tenure at the Oregon State University Department of Mathematics (1961–80), Kaplan never published in scientific journals. After many years of silence, in 1982 he published his tutorial book on *Mathematical Programming and Games*.

In a sense, the product–limit estimator of the Kaplan–Meier survival curve was a simple version of more complex models for state transitions, as he had already discussed in his doctoral thesis of 1950 (Kaplan 1950b, 1953, 1955a). In the 1982 book he took analyses of conditional processes to a mathematically higher level. John Nash, whose works and mathematical solutions are amply cited in Kaplan's book, commented:

> The book is of the type of books on Game Theory that emphasize very much the connections of 'Game Theory' and 'Linear Programming'. These books are, of course, good for you to have if you are interested in those two theoretical topics and their interrelations
>
> (J R Nash, Jr, personal communication to János Kollár, 30 June 2014).

In relation to game theory, the main topic of Nash's own research, Nash added:

> ... but otherwise if you are more broadly interested in Game Theory then you find that the coverage of games for themselves is not so broad.

During our research after Kaplan's death, we also discovered a full manuscript describing a new notation system for musical chords (Kaplan 1998).

Kaplan had tried in vain to get the paper published; it was rejected by both musicology journals and by mathematical journals. Despite the several versions of the manuscript which we discovered together with manuscript submission letters, we did not find a complete response from an editor, nor did we locate a referee's report.

The fate of the other inventors

The 1958 Kaplan and Meier paper greatly facilitated survival analysis: both in medicine and for industrial applications. In fact, Kaplan had originally developed the mathematical method to estimate the mean lifetime of light bulbs when working at Bell Laboratories (Kaplan 1955a; Garfield 1983). Many extensions and refinements have since made survival analysis a powerful tool; the Kaplan–Meier method is still the mathematical basis for all these methods. The distinguished British epidemiologist Sir Richard Peto said 'I and hundreds of others use his methods every week in our work' (Washington Post, quoted in Brown 2011).

Princeton Professor Samuel S Wilks died suddenly at the age of 57, in 1964. He was remembered as 'The Statesman of Statistics', for his outstanding work in multivariate analysis, as an inspiring educator and promotor, and as an advisor to government agencies and other research organizations. The American Statistical Association (ASA) named its most prestigious annual award after him 'recognizing outstanding contributions to statistics that carry on in the spirit of his work' (Mosteller 1964; Schultz and Eisenberg 2000; ASA 2004).

At Princeton, Wilks had delegated the daily PhD supervision of both Kaplan and Meier to the youthful professor John W Tukey (1915–2000). In 1965, with James Cooley, Tukey had introduced an analytical tool known as fast Fourier transform, which remains a ubiquitous technique for understanding waveforms in fields from astrophysics to electrical engineering (Schultz and Eisenberg 2000; Brillinger 2002). Among many honours and awards, Tukey received the National Medal of Science in 1973 and an honorary doctorate from Princeton in 1998, and was a member of the National Academy of Sciences and the Royal Society. In 1965, he received the second Samuel S Wilks award of the ASA. The biography of Tukey published in the same year mentions Paul Meier—but not Kaplan—in his list of fifty-five PhD students (Dressel *et al.* 1965; Brillinger 2002). Paul Meier is also listed by the American Mathematical Society as one of Tukey's PhD students in the mathematics genealogy project (AMS 2016b); again, Kaplan is not listed. Kaplan's absence from these lists remains a mystery.

Paul Meier (1924–2011) published his Princeton PhD thesis in 1951. Next, he taught biostatistics at John Hopkins University, and in 1957 joined the Statistics Department at the University of Chicago. He remained there until 1992, and spent the last years of his career at Columbia University (Marks 2004). Amongst many awards, in 2004 he received the 41st Samuel S Wilks award of the ASA, for his important work 'including the Kaplan–Meier estimator for survival analysis' (ASA 2004). Edward Kaplan did not share in the award, was not invited to the ceremony and was not notified about it.

The eclipse of Edward Kaplan

Meier held several prestigious advisory positions, and his death at age 87 was memorialized in international journals and newspapers, including obituaries in the New York Times, the Chicago Tribune, Washington Post, The Lancet and by the American Mathematical Society (Brown 2011; Hevesi 2011; Pincock 2011; Tobacman 2011).

By contrast, the death of Edward Kaplan in 2006 was announced in an obituary in the local newspaper in Corvallis, Oregon and also in the newsletter of the Institute for Mathematical Statistics (Anonymous 2006; Smythe 2006). The American Mathematical Society included a death notice of two and a half lines in a list of forty recently deceased AMS members: 'Edward L. Kaplan, from Corvallis, OR, died on September 26, 2006. Born on May 11, 1920, he was a member of the Society for 59 years' (AMS 2007).

Thus, although Kaplan was a longstanding member of the major US mathematical and statistical societies, the death of their most cited member went virtually unnoticed. Kaplan never received a medal or award. Neither his name nor his portrait can be found in the annals or Hall of Fame at Carnegie Mellon, Princeton, Bell Labs, Berkeley or Corvallis. There is no Edward L Kaplan library, lecturer, fellowship, concert hall, university building nor coffee corner. In short, despite his unquestionable role in the 1958 paper, he was given little or no credit nationally or internationally. There is even some evidence to suggest that his colleagues at Oregon State University were not aware of his role in the formulation of the KM methodology.

What happened to Edward L Kaplan and why was he seemingly not recognized? Was he simply overlooked, or were there other reasons?

His detailed handwritten journals, which he kept from the age of eleven, indicate that he excelled academically from early childhood through his high school graduation (as valedictorian) in 1937. His social life was centred around his family: especially his brother, Don, and his parents in Pennsylvania. As a student in Pittsburgh, as we have seen, his scholastic accomplishments were consistent, and consistently recognized in the Putnam Awards; and he spent the next twenty years in esteemed academic institutes. His output included about sixty mostly classified scientific manuscripts, his 1950 PhD thesis and of course the paper of 1958. Such a combination of intelligence, commitment and opportunity were later described by the highly respected mathematician and computer scientist, Richard Hamming (1915–98), as a major ingredient for success in science (Hamming 1986). During Kaplan's years at Bell Labs, Richard Hamming was his friend and best man at his wedding to Frances Berting in 1958. Hamming mentioned another ingredient of success which he considered 'very distasteful' but indispensable:

> it is not sufficient to do a job, you have to sell it. Selling to a scientist is an awkward thing to do. It's very ugly; you shouldn't have to do it. The world is supposed to be waiting, and when you do something great, they should rush out and welcome it
>
> (Hamming 1986).

From our intensive investigation of several sources including colleagues at work and his family, and his students, it seems that Kaplan did not 'sell' himself or his work during his training and subsequent career. On the contrary, he left research for a teaching position in the Department of Mathematics at Oregon State University: both geographically and intellectually, he distanced himself from the network of researchers and friends who knew the value of his work and could promote his scientific accomplishments. It was also very clear that, for whatever reason, he did not wish to be remembered for his previous work. He never cited his own papers, and neither did his PhD students and master's students in their theses; there is thus some doubt whether they were aware of his previous papers. From his personal notebooks and from our interviews with his students, former colleagues and family, it is also clear that over the years he increasingly distanced himself from his mathematics colleagues both academically and socially.

The roots of this behaviour are perhaps to be sought in a series of personal crises beginning with the loss of his brother.

> My only sibling Donald Eugene Kaplan was born May 20, 1923, after our parents moved from the Philadelphia area to Swissvale, PA. My first firm memory is his being nursed, and my being given a taste. He was three years younger than I, and four years behind me in school; thus we both graduated (from high school and Carnegie Tech) in June 1941. Physically and socially Don was my equal or superior; thus we were rivals to some extent, but became very close when we last met before he became a wartime casualty in France on November 18, 1944
>
> (Kaplan 1999, 2–3) (see Figure 5).

Edward Kaplan himself did not join the army, since his parents wanted, so to speak, 'to save the brains of the family'. In 1941, he verbalized his engagement with war and peace in a manuscript 'The Attainment of International Peace and Justice',

Figure 5. Donald and Edward Kaplan on horseback, 1940 (Source with permission: Kaplan family archive)

a manuscript that is unfortunately lost (Kaplan 1999). In his last letter to Edward, Donald wrote that he realized his odds of returning alive from war-ravaged Europe were only 2 in 10. Donald was said to have been killed while leaving his shelter to save a wounded fellow soldier.

The death of his brother haunted Edward Kaplan for the rest of his life. After the war, he appeared to change from an extroverted young man to an increasingly reclusive person. There were periods of anxiety accompanied by tension problems. Kaplan respected the symbolic funeral of Donald at Arlington National Cemetery in 1948, but realized that there were no remains under the cross at the burial site, and that his brother was definitively lost to follow-up. 'The world was so wicked', he wrote in his diary and he decided never to have children.

In a later collection of his own poems entitled 'Invitation to Mortality', Edward Kaplan (1985) wrote:

> My God, why do I still linger?
> How many times can one spirit
> Die and yet rise again?

Kaplan's textbook on *Mathematical Programming and Games* explains some of the mathematical ways in which he tried to deal with the psychological turmoil of this period (Kaplan 1982). On page 166, Kaplan took up the Prisoner's Dilemma, for which John Nash formulated a mathematical solution, the so-called Nash equilibrium (Nash 1951, 1953). The original dilemma comes from game theory: two prisoners each have to choose between being 'altruistic' and 'egoistic'; the penalty depends on the other prisoner's choice. Kaplan developed the problem in a chapter called 'The Prisoner's Trilemma', identifying three possible choices: altruistic, pragmatic or psychopathic.

> The pragmatist expects the other player to be like himself, willing to accept a small loss rather than slaughter the other player, but also willing to take what he can on a one-for-one basis. ('Your loss is no more than my gain') A sadistic or nihilistic strategy with nonpositive payoffs could be added to complete the psychological spectrum
>
> (Kaplan 1982, 166).

Perhaps we should read 'soldier' for 'prisoner'. Kaplan's papers, textbooks and poems are full of such allusive passages, dealing with lost loved ones, unattainable loves, and reconciliation with mortality and survival (Kaplan 1985, 1999).

A second crisis, closer in time to the period of Kaplan's crucial failure to 'sell' himself as a mathematician, was his 1961 divorce. During his time at the Bell Laboratories, the period when he started working on what was to become the Kaplan–Meier survival curve and many other projects, he met Frances Berting, a materials engineer. They were married in 1957 and moved to California, but the marriage ended only four years later. Dr Berting recalled that Kaplan had never told her anything about his work with Paul Meier. Similarly at Livermore, since his work was classified, he never told her what he precisely was working on. She could confirm the family gossip that during that time 'he was surrounded by the grey suits', indicating that he was frequently accompanied by FBI agents which he found more annoying than threatening: 'O yes, those days at Livermore, it was not unusual to be surrounded by security guys. Most of the work was classified work for the war machinery'.

Dr Berting has further confirmed the reports of Kaplan's own personal notes that during that period, and maybe before, he had periods of seemingly mentally paralysing anxiety, and took extended sick leaves from Bell Labs and Berkeley Livermore. Discontent about his work for the development of the hydrogen bomb made for a third source of personal crisis.

> My stint of teaching in 1960 was enjoyable, but the talk of hydrogen bombs was disconcerting. Hence in 1961 I resolved to look for a teaching position not within a metropolitan area
>
> (Kaplan 1999).

Kaplan quit his job at the Lawrence Livermore Lab and, without saying farewell or even leaving a forwarding address, moved alone to Corvallis, Oregon, to teach mathematics at Oregon State University as a professor of mathematics. By this decision, his career turned from research to an emphasis on teaching mathematics. From that time on, his mathematical interests were almost solely devoted to teaching and his academic mathematical output decreased, although at the same time he remained actively involved with music and with literature.

Teaching under less strenuous conditions did indeed give him more balance in life, but he became increasingly shy. Those who knew him personally remember him as a gifted educator: smart, with some dry wit. His graduate student, Wally Reed, remembered that he gave the most lucid lectures on the most complex mathematical problems, yet:

> All the notes he used for a lecture were written on the inside of a paper match folder
>
> (Wally Reed, personal communication, 2015).

His PhD student Dr Donald Cresswell remembers:

> He had a brilliant mind, but at first encounter quite intimidating. He was not really sociable with the other professors, but one-on-one he was a kind and generous person. He had a huge grand-piano in his apartment, and had several lady-friends with whom he enjoyed dancing. One might say he lived a solitary life. I corresponded with him every Christmas, and only got notice of his death when my last Christmas card was returned to sender.
>
> (Donald Cresswell, personal communication, 2015)

After the death of his parents in 1978, Edward Kaplan seemingly felt increasingly uncomfortable even in teaching. He retired early in 1980 (see Figure 6).

In later life Kaplan considered the family of his cousin Vernon and his wife Eunice, in Owatonna, Minnesota, a second home. His Minnesota family recalled that he maintained his interests, including music; a cousin, Jeanette (Meixner) Franzel, remembers:

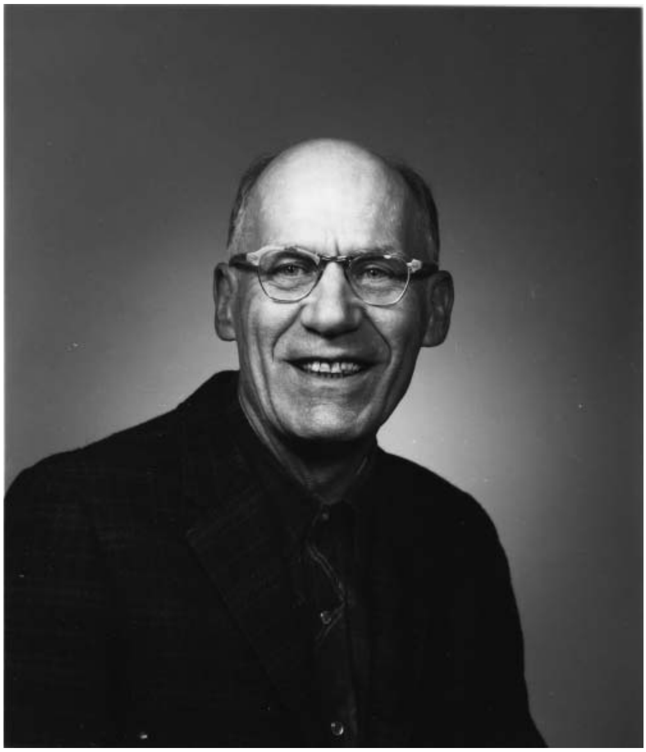

Figure 6. Edward L Kaplan in 1981. (Source with permission: Oregon State University)

> I will never forget when he came the year the Rubic's Cube puzzle was popular, he sat in my dad's reading chair in our 'oak room' and just stared at the toy and working to solve it mathematically without making any moves
>
> (Jeanette Franzel, personal communication, 2015).

In 1994 Kaplan met Mae Nichol, who became his second wife, and to whom he transferred his love for music, opera and theatre, and who cared for him in his later days when he was medically frail. She related that:

> He was such a nice and refreshing person. So humble. Edward never bragged about his famous work, and he showed no resentment for the lack of praise. Very few probably knew that he did ground breaking work. He was shy in public, but he was very sociable amongst friends and family. A listener more than a talker. He played the piano very well, he wrote poems and composed and recorded some beautiful music
>
> (Mae Nichol-Kaplan, personal communication, 2015).

During the latter years of his life these interests and activities declined due to his deteriorating medical condition. In 1999, long after his retirement, he was diagnosed with normal pressure hydrocephalus, a debilitating neurological disease which—with hindsight—may have been affecting his mental abilities and movement for many years previously. A neurosurgical intervention that year did not result in significant relief (according to him, and also confirmed to us by his physician). After a long and debilitating illness, Edward Kaplan died on 26 September 2006 as a result of superimposed pneumonia.

His obituary in the Corvallis Gazette-Times, Oregon mentioned that 'by request of Edward, there will be no services, but his memory will live on' (Anonymous 2006). His ashes are buried next to his parents in Owatonna, Minnesota, in the National Bohemian Cemetery, not far from the farm of his grandfather, Joseph, who emigrated in the mid-nineteenth century from Bohemia to the United States.

The Kaplan–Meier method deals with life, death and those 'lost to follow-up'. For too long, Edward L Kaplan has been undeservingly lost to follow-up.

Acknowledgements

Mrs Mae Nichol-Kaplan from Corvallis, Oregon, and Mrs Eunice Meixner from Owatonna, MN, and her children Jeanette, Tom, Ken, Karen, Sheila Andy and Mary Meixner, gave us a wealth of personal information. We greatly thank them for their immeasurable help and hospitality while these data were being collected. Mrs Eunice Meixner gathered an amazing collection of records and artefacts relating to Edward Kaplan and she made all of these available for our investigation. Dr Frances Berting, Edward Kaplan's first wife, was also extremely helpful to us as we organized information about those years of his life.

Dr Donald J Cresswell, Dr Wally Reed, Dr Norman Franzen, Professor Ronald Guenther, Professor Diane Meier, and Mrs Karen Meier, shared many valuable personal memories.

Mr Geoff M Somnitz, reference student, and Mr Michael A Dicianna of the Special Collections and Archives Research Center, Professor Sastry G Pantula, Dean of the College of Science and Robert T Smythe, PhD, Professor Emeritus of statistics all from Oregon State University, Corvallis, Oregon, provided valuable information from the Oregon State University archives. We are very grateful to them.

Dr Jennifer Johnson, and Professor János Kollár, mathematicians at Princeton University, and Edward J. Eckert, and Ms Rebecca Nadolny, archivists of Alcatel-Lucent at Bell Labs, Murray Hill NJ, were helpful in locating and making available early manuscripts and reports written by Edward Kaplan. The late Professor John R Nash of Princeton University was so kind as to comment about Kaplan's 1982 book.

Professor Edmund A Gehan, statistician, MD Anderson Cancer Center, Houston TX, provided valuable information on the history of survival analysis.

Ms Amy Claussen and Mr Del Reed, research librarians of the Biomedical Library of the University of Minnesota Medical School in Minneapolis, assisted us in collecting the data for the citation analysis of the 1958 KM published manuscript.

Special thanks for the assistance of Mr Alex Schueller from the Office of then United States Senator Al Franken (Minnesota) in obtaining personnel records (with permission of Edward Kaplan's widow); with Mr Schueller's assistance we also obtained access to records from the Naval Ordnance Laboratory. Copies of Kaplan's personnel file from the Naval Ordnance Laboratory were provided by Ms Cara Moore and Ms Ashley Mattingly, archivists of the United States National Archives in St Louis, Missouri. (Written permission was obtained from Professor Kaplan's widow.)

Professor Sylvia Nasar, Columbia University, New York, New York provided us with daily journal notes from Edward Kaplan which had been previously sent to her by Ms Mae Nichol. This was an extremely valuable group of documents.

Dr Richard Lafrance, MD, retired neurologist at Corvallis Clinic, after written consent from Mae Nichol-Kaplan, reviewed the medical history of Kaplan during the course of the normal pressure hydrocephalus. This proved very valuable for evaluating the last decade of his life.

Mr Lars Groot, medical student at the AMC in Amsterdam, The Netherlands, arranged the fifty Halsted patients into a life-table for further analysis.

The authors are grateful to Claus Pierach, MD for his evaluation of the manuscript and his insightful suggestions. We also express our gratitude to Mrs Irene S Kaplan, the wife of co-author Edward L Kaplan, MD, for her very perceptive suggestions during our investigations and writing for this manuscript.

NOTE: The co-author of this manuscript, Edward Lawrence Kaplan, MD, Professor Emeritus, Department of Pediatrics, University of Minnesota Medical School, Minneapolis, Minnesota, is not related to the subject of this manuscript, Edward Lynn Kaplan, PhD.

Disclosure statement

No potential conflict of interest was reported by the author.

Supplemental data

Supplemental data for this article can be accessed https://doi.org/10.1080/17498430.2018.1450055.

ORCID

Lukas J A Stalpers http://orcid.org/0000-0001-9574-0295

References

AMA—American Mathematical Association, *Putnam Competition Individual and Team Winners*, 2016a, http://www.maa.org/programs/maa-awards/putnam-competition-individual-and-team-winners (accessed on 29 November 2017).

AMS—American Mathematical Society, *Mathematical Genealogy Project*, 2016b, http://www.genealogy.ams.org/id.php?id=136177 (accessed on 29 November 2017).

AMS—American Mathematical Society, 'Deaths of AMS Members', *Notices of the AMS*, 54(4) (2007), 544, http://www.ams.org/notices/200704/inside-web.pdf (accessed on 29 November 2017).

Anonymous, 'Obituary on Edward L. Kaplan', *Corvallis Gazette-Times*, 27 September 2006, http://www.gazettetimes.com/news/local/obituaries/edward-kaplan/article_3abb619d-834f-5e19-b34d-3b54ccfd9cdb.html (accessed on 29 November 2017).

Armitage, P, 'The comparison of survival curves', *Journal of the Royal Statistical Society, Series A-general*, 122(3) (1959), 279–300.

ASA—American Statistical Association. 'Samuel S. Wilks Award. 2004 Paul Meier', 2004, http://www.amstat.org/ASA/Your-Career/Awards/Samuel-S-Wilks-Memorial-Award.aspx (accessed on 29 November 2017).

Berkson, J, 'Estimation of the interval rate in actuarial calculations: a critique of the person-years concept (Summary)', *Journal of the American Statistical Association*, 49 (1954), 363.

Berkson, J, and Gage, R P, 'Calculation of survival rates for cancer', *Proceedings of the Staff Meetings of the Mayo Clinic*, 25 (1950), 270–286.

Berkson, J, and Gage, R P, 'Survival curve for cancer patients following treatment', *Journal of the American Statistical Association*, 47 (1952), 501–515.

Bodey, G P; Gehan, E A; Freireich, E J, and Frei, E, 3rd,'Protected environment-prophylactic antibiotic program in the chemotherapy of acute leukemia', *American Journal of Science*, 262(3) (1971), 138–151.

Borkowf, C B, 'A simple hybrid variance estimator for the Kaplan–Meier survival function', *Statistics in Medicine*, 24 (2005), 827–851.

Brillinger, D R, 'John W. Tukey: his life and professional contributions', *The Annals of Statistics*, 30(6) (2002), 1535–1575.

Brown, D, 'Paul Meier, biostatistician and co-inventor of a famous graph, dies at 87', *Washington Post*, 11 August 2011, http://www.washingtonpost.com/local/obituaries/paul-meier-biostatistician-and-co-inventor-of-a-famous-graph-dies-at-87/2011/08/09/gIQAfHBH7I_story.html (accessed on 29 November 2017).

Cho, H; Howlader, N; Mariotto, A B, and Cronin, K A, 'Estimating relative survival for cancer patients from the SEER Program using expected rates based on Ederer I versus Ederer II method', *Surveillance Research Program, NCI, Technical Report #2011-01*, Bethesda, MD, 2011 (accessed on 29 November 2017).

Cole, S R, and Hudgens, M G, 'Survival analysis in infectious disease research: describing events in time', *AIDS*, 24(16) (2010), 2423–2431.

Coltman, C A, Jr; Costanzi, J J; Dudley, G M, 3rd; Haut, A; Lane M, and Gehan, E A, 'Further clinical studies of combination chemotherapy using cyclophosphamide, vincristine, methotrexate and 5-flourouracil in solid tumors', *The American Journal of the Medical Sciences*, 261(2) (1971), 73–78.

Cox, D R, 'The analysis of exponentially distributed lifetimes with 2 types of failure', *Journal of the Royal Statistical Society, Series B*, 21(2) (1959), 411–421.

Cox, D R, 'Regression models and life tables', *Journal of the Royal Statistical Society, Series B*, 34 (1972), 187–220.

Cutler, S J, and Ederer F, 'Maximum utilization of the life table method in analyzing survival', *Journal of Chronic Diseases*, 8 (1958), 699–712.

Dickman, P W, and Coviello, E, 'Estimating and modeling relative survival using SAS', *The Stata Journal*, 15(1) (2015), 186–215, http://www.pauldickman.com/survival/strs.pdf (accessed on 29 November 2017).

Doob, J L, 'Renewal theory from the point of view of the theory of probability', *Transactions of the American Mathematical Society*, 63 (1948), 422–438.

Doob, J L, *Stochastic processes*, New York: John Wiley and Sons, 1953 [The original manuscript used by Kaplan was planned as a book by J L Doob and N Wiener: Time Series Analysis].

Dressel, F G; Eisenhart, C; Kempthorne, O; Mood, A M; Simon, L E, and Grubbs, F E, 'Tukey awarded Wilks Memorial Medal', *The American Statistician*, 19(5) (1965), 2, http://www.tandfonline.com/doi/abs/10.1080/00031305.1965.10479752?journalCode=utas20 (accessed on 29 November 2017).

Ederer, F, 'A parametric estimate of standard error of survival rate', *Journal of the American Statistical Association*, 56(293) (1961a), 411–421.

Ederer, F; Axtell, L M, and Cutler, S J, 'The relative survival rate: a statistical methodology', *National Cancer Institute Monographs*, 6 (1961b), 101–121.

Ederer, F, and Heise, H, *Instructions to IBM 650 programmers in processing survival computations* Technical, End Results Evaluation Section, National Cancer Institute, Bethesda, MD, 1959.

Easton, J, 'The curve that changed the world', *Science Life*, University of Chicago, 10 August 2011, http://sciencelife.uchospitals.edu/2011/08/10/the-curve-that-changed-the-world/ (accessed on 29 November 2017).

Fine, J P, and Gray, R J, 'A proportional hazards model for the subdistribution of a competing risk', *Journal of the American Statistical Association*, 94(446) (1999), 496–509.

Fuller, L M; Gamble, J F; Shullenberger, C C; Butler, J J, and Gehan, E A, 'Prognostic factors in localized Hodgkin's disease treated with regional radiation. Clinical presentation and specific histology', *Radiology*, 98(3) (1971), 641–654.

Garfield, E, 'Interview with Sidney J. Cutler in 'This Week's Citation Classic' on: Cutler SJ and Ederer F. Maximum utilization of the life table method in analyzing survival. Journal of Chronic Diseases, 8:699–712, 1958', *Current Contents*, 16 (1979), 356.

Garfield, E, 'This week's citation classic. Interview on the Kaplan and Meier, 1958 publication', *Current Contents*, 24 (1983), 14.

Garfield E, 'The most cited papers of all time, SCI 1945–1988. Part 1A. The SCI top 100 – Will the Lowry method ever be obliterated?', *Current Contents*, 7 (1990), 3–14.

Gehan, E A, 'A generalized Wilcoxon test for comparing arbitrarily singly censored samples', *Biometrika*, 52 (1965), 203–223.

Gehan, E A, 'Estimating survival functions from life table', *Journal of Chronic Diseases*, 21(9–10) (1969), 629–644.

Gilfillan, E S, and Kaplan, E L, 'Transient torques in squirrel-cage induction motors, with special reference to plugging', *Transactions of the American Institute of Electrical Engineers*, 60(12) (1941), 1200–1209.

Green, S B, 'A conversation with Fred Ederer', *Statistical Science*, 12(2) (1997), 125–131.

Greenwood, M, 'The natural duration of cancer', *Reports on Public Health and Medical Subjects*, No 33, London: His Majesty's Stationery Office, 1926.

Halsted, W S, 'The results of operations for the cure of cancer of the breast performed at the Johns Hopkins Hospital from June, 1889, to January, 1894', *Annals of Surgery*, 20 (1894), 497–555.

Hayward, T E, 'On life-tables – their construction and practical application (parts I and II)', *Journal of the Royal Statistical Society*, 62(3) (1899a), 443–488.

Hayward, T E, 'On life-tables: their construction and practical application (part III)', *Journal of the Royal Statistical Society*, 62(4) (1899b), 683–702.

Hayward, T E, 'On life-tables: their construction and practical application (addendum)', *Journal of the Royal Statistical Society*, 63(4) (1900), 625–636.

Hayward, T E, 'Notes on life-tables', *Journal of the Royal Statistical Society*, 65(2) (1902a), 354–358.

Hayward, T E, 'Some additional notes on a 'short' method of constructing life-tables', *Journal of the Royal Statistical Society*, 65(4) (1902b), 680–684.

Hamming, R W, 'You and your research', *Transcription of the Bell Communication Research Colloquium Seminar*, 7 March 1986, http://www.cs.virginia.edu/robins/YouAndYourResearch.pdf and https://www.youtube.com/watch?v=a1zDuOPkMSw (last accessed 29 November 2017).

Hevesi, D, 'Paul Meier, Statistician Who Revolutionized Medical Trials, Dies At 87', *New York Times*, 12 August 2011, http://www.nytimes.com/2011/08/13/health/13meier.html (last accessed 29 November 2017).

Howlader, N; Noone, A M; Krapcho, M; Miller, D; Bishop, K; Altekruse, S F; Kosary, C L; Yu, M; Ruhl, J; Tatalovich, Z; Mariotto, A; Lewis, D R; Chen, H S; Feuer, E J, and Cronin, K A (eds), *SEER Cancer Statistics Review 1975–2014* (released 14 April 2017), National Cancer Institute, Surveillance, Epidemiology, and End Results Program, Bethesda, MD, 2016, https://seer.cancer.gov/csr/1975_2014/ (accessed 4 March 2018).

Kaplan, E L, 'Properties of underwater diaphragms', *NOLM 1629*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 1 June 1942a.

Kaplan, E L, 'Magnetic fields of square coils, singly and in pairs', *NOLR 574*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 16 June 1942b.

Kaplan, E L, 'The vertical vs. the horizontal position for the Mk 13 mine with M-4 Mod 1 mine mechanism', *NOLM 2032*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 3 August 1942c.

Kaplan, E L, 'The fields of circular coils', *NOLR 785*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 12 May 1943a.

Kaplan, E L, 'Performance of the MK 16 mine', *NOLM 3749*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 27 May 1943b.

Kaplan, E L, 'Comparison of effectiveness of four types of depth charge', *NOLM 3852*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 14 June 1943c.

Kaplan, E L, 'The complete field of a circular coil', *NOLR 825* Naval Ordnance Laboratory, Navy Yard, Washington DC, 11 August 1943d.

Kaplan, E L, 'Predicted response of the gradiometer-operated XFDC in the observed fields of submarines', *NOLM 4247*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 8 September 1943e.
Kaplan, E L, 'Predicted target areas for the gradiometer-operated XFDC', *NOLM 4341*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 30 September 1943f.
Kaplan, E L, 'Analysis of timing-period data for sterilizers SD-4 and 5', *NOLM 4353*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 2 October 1943g.
Kaplan, E L, 'Predicted performance of a magnetically-fuzed mousetrap projectile', *NOLM 4370*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 13 October 1943h.
Kaplan, E L, 'Typical XTFDC signatures including parasitic signal', *NOLM 4425*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 19 October 1943i.
Kaplan, E L, 'Comparison of predicted positional firing characteristics of gradimeter and total-field proximity depth charges', *NOLM 4614*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 27 November 1943j.
Kaplan, E L, 'Computed trajectories of parachute mines', *NOLM 4672*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 10 December 1943k.
Kaplan, E L, 'Magnetic properties of permalloy cores' (editorial assignment only), *NOLR 565*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 26 January 1944a.
Kaplan, E L, 'Predicted performance of the A-3 mechanism', *NOLM 5041*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 25 February 1944b.
Kaplan, E L, 'Actuation tests of XLF-12a mechanism with recorded ship signatures', *NOLM 5505*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 24 August, 1944c.
Kaplan, E L, 'General table of mine dip and drift', *NOLM 6558*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 2 January 1945a.
Kaplan, E L, 'Actuation tests of A-5 mechanism with recorded ship signatures', *NOLM 6973*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 24 March 1945b.
Kaplan, E L, 'Oscillating mines Mk 19 and XOM-theoretical study', *NOLM 7119*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 17 April 1945c.
Kaplan, E L, 'Water-entry velocity of parachute mines', *NOLM 7503*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 28 June 1945d.
Kaplan, E L, 'Calculated effectiveness of depth charges Mk 8 and 11 as influenced by submarine speed and type of sound gear', *NOLM 8225*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 27 December 1945e.
Kaplan, E L, 'Predicted performance of total-field mechanism in DC cases Mk 8 and 11', *NOLM 8226*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 27 December 1945f.
Kaplan, E L, 'Auxiliary table of complete elliptic integrals', *Journal of Mathematical Physics*, 25 (1946a), 26–36 (also published under: RMT 311, MTAC-Mathematical Tables and other Aids to Computation, 311(2) (1946–1947), 127–128).
Kaplan, E L, 'Mathematical comparison of two gun-interference eliminators', *NOLM 8905*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 12 December 1946b.
Kaplan, E L, 'Classification of total-field signatures of submarines along vertical lines', *NOLM 8884*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 16 January 1947a.
Kaplan, E L, 'Mathematical analysis of intermittent arming of mines', *NOLM 8923*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 23 January 1947b.
Kaplan, E L, 'Mathematical comparison of delayed and intermittent arming of mines', *NOLM 9115*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 22 May 1947c.
Kaplan, E L, 'Statistical analysis of vibration tests on M-13 mechanisms', *NOLM 9332*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 9 December 1947d.
Kaplan, E L, 'The sources of shooting errors, and their dependence on type of weapon and method testing', *NOLM 9452*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 8 December 1947e.

Kaplan, E L, 'Comparison of estimated cumulative kill probabilities for arrowshells with and without guidance, a ramjet-assisted projectile, and conventional 5" and 8" projectiles', *NOLM 9453*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 8 December 1947f.

Kaplan, E L, 'Auxiliary table for the incomplete elliptic integrals', *Journal of Mathematical Physics*, 27 (1948a), 11–36.

Kaplan, E L, 'One-shot kill probabilities for guided arrowshells and other projectiles', *NOLM 9559* (similar to NOLM 9662), Naval Ordnance Laboratory, Navy Yard, Washington DC, 3 February 1948b.

Kaplan, E L, 'The sequential t-test', *NOLM 9613*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 7 April 1948c.

Kaplan, E L, 'Design calculations for 6"/47 illuminating projectiles Mk 38 and 41, with F-type contents', *NOLM 9333*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 1948d.

Kaplan, E L, 'Sono-optical recorder: mathematical study', *NOLM 9634*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 1948e.

Kaplan, E L, 'One-shot kill probabilities as influenced by antiaircraft system parameters', *NOLR 9662*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 1 June 1948f.

Kaplan, E L, 'Mathematical developments relating to aircraft kill probabilities', *NOLR 1087*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 1 October 1948g.

Kaplan, E L, 'Multiple elliptic integrals', *Journal of Mathematical Physics*, 29 (1950a), 69–75.

Kaplan, E L, 'Infinite permutations of stationary random sequences', Doctoral dissertation, Princeton University, Princeton, NJ, 1950b (revised and extended version in 1953).

Kaplan, E L, 'Tensor notation and sampling cumulants of K-statistics', *Biometrika*, 39(3–4) (1952a), 319–323.

Kaplan, E L, 'Numerical integration near a singularity', *Journal of Mathematical Physics*, 31 (1952b), 1–28.

Kaplan, E L, 'The reduction of incomplete life test data', Bell Laboratories Internal Report (dn TM53-1400-38 / da 11/10/1953), Murray Hill, NJ, 1953.

Kaplan, E L, 'Transformation of stationary random sequences', *Mathematica Scandinavica*, 3 (1955a), 127–149.

Kaplan, E L, 'Statistical analysis of incomplete rectangular tables', *Bell Laboratories Internal Report* (da 6/01/1955. dn TM55-114-27; TM55-114-27; tm55-114-27), Murray Hill, NJ, 1955b.

Kaplan, E L, 'Signal-detection studies, with applications', *Bell System Technical Journal*, 34(2) (1955c), 403–437.

Kaplan, E L, '50 and 95%-points for sums of squares for two or three Gaussian variables', *Bell Laboratories Internal Report* (da 12/31/1956. dn TM56-114-43), Murray Hill, NJ, 1956.

Kaplan, E L, 'Monte Carlo methods for equilibrium solutions in neutron multiplication', *U.S. Atomic Energy Commission*, University of California Lawrence Radiation Laboratory, Livermore, CA (available from Office of Technical Services, US Department of Commerce), 1958a.

Kaplan E L, 'Monte Carlo methods', *Proceedings of the Fifth Annual High-Speed Computer Conference*, Louisiana State University, Baton Rouge, LA, 1958b.

Kaplan, E L, 'Backward Monte Carlo and other methods for 'reversing' stochastic processes', *UCRL 6649*, University of California, Radiation Laboratory, Livermore, CA, 1961a.

Kaplan, E L, 'Isotropic constant-cross-section calculations on neutron multiplication in a small bare homogeneous sphere', *UCRL 6660*, University of California, Radiation Laboratory, Livermore, CA, 1961b.

Kaplan, E L, *Application letter of 1961*, Oregon State University Library/Special Collections and Archives Research Center, Corvallis, Oregon, 1961c.

Kaplan, E L, *Mathematical programming and games. Volume I*, New York: John Wiley and Sons, 588 pages, 1982.

Kaplan, E L, *Invitation to mortality* (a collection of unpublished poems), Corvallis, OR, 77 pages (limited number self-publication), 1985.

Kaplan, E L, *Chord and set symbols based on the 352 types*, unpublished manuscript, 1998.

Kaplan, E L, *The life and reflections of Edward L. Kaplan* (autobiographical notes), Corvallis, OR (limited number self-publication), 1999.

Kaplan, E L, and Meier, P, 'Nonparametric estimation from incomplete observations', *Journal of the American Statistical Association*, 53(282) (1958), 457–481.

Kaplan, E L, and Mooney, R L, 'Theoretical study of subsonic radiation under water', *NOLM 4016*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 21 July 1943.

Kaplan, E L, and Slawsky Z I, 'Comparison of estimated rounds per bird for arrowshells with and without guidance, a ramjet-assisted projectile, and conventional 5" and 8" projectiles', *NOLM 9331*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 2 October 1947d.

Kaplan, E L, and Terry, M E, 'The adjustment of merit ratings to eliminate rater biases', *Bell Laboratories Internal Report* (dn TM54-114-16 / da 03/01/1954), Murray Hill, NJ, 1954.

Kaplan, E L, and Yost, F L, 'Average firing efficiency of two M-13 mechanisms with weapon A sinking rate', *NOLM 9256*, Naval Ordnance Laboratory, Navy Yard, Washington DC, 19 April 1948.

Keynes, G, 'The treatment of primary carcinoma of the breast with radium', *Acta Radiologica*, 10 (1929), 893–901.

Korteweg, J A, 'Carcinoom en statistiek (Carcinoma and statistics)', *Nederlands Tijdschrift voor Geneeskunde*, 47 (1903), 1054–1068.

Lathan, S R, 'Caroline Hampton Halsted: the first to use rubber gloves in the operating room', *Baylor University Medical Center Proceedings*, 23(4) (2010), 389–392.

Lin, D Y, and Wei, L J, 'The robust inference for the Cox proportional hazards model', *Journal of the American Statistical Association*, 84(408) (1989), 1074–1078.

Luce, J K; Gamble, J F; Wilson, H E; Monto, R W; Isaacs, B L; Palmer, R L; Coltman, C A, Jr; Hewlett, J S; Gehan, E A, and Frei, E, 3rd, 'Combined cyclophosphamide vincristine, and prednisone therapy of malignant lymphoma', *Cancer*, 28(2) (1971), 306–317.

Marks, H M, 'A conversation with Paul Meier', *Clinical Trials*, 1 (2004), 131–138.

May, S, and McKnight, B, 'Graphics and statistics for cardiology: survival analysis', *Heart*, 103 (2017), 335–340.

Meier, P, *Weighted means and lattice designs*, Doctoral thesis, Princeton University, Princeton, NJ, 1951.

Meier, P, 'Variance of a weighted mean', *Biometrics*, 9(1) (1953), 59–73.

Miettinen, O S, 'Survival analysis: up from Kaplan–Meier–Greenwood', *European Journal of Epidemiology*, 23(9) (2008), 585–592.

Mosteller, F, 'Samuel S. Wilks: statesman of statistics', *The American Statistician*, 18(2) (1964), 11–17.

Nash, J, 'Non-cooperative games', *Annals of Mathematics*, 54 (1951), 286–295.

Nash, J, 'Two-person cooperative games', *Econometrica*, 21 (1953), 128–140.

Peto, J, 'The calculation and interpretation of survival curves', in M E Buyse, M J Staquet, and R J Sylvester (eds), *Cancer clinical trials. Methods and practice*, Oxford: Oxford Medical Publications, Oxford University Press, 1984, 361–380.

Peto, R; Pike, M C; Armitage, P; Breslow, N E; Cox, D R; Howard, S V; Mantel, N; McPherson, K; Peto, J, and Smith, P G, 'Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design', *British Journal of Cancer*, 34(6) (1976), 585–612.

Peto, R; Pike, M C; Armitage, P; Breslow, N E; Cox, D R; Howard, S V; Mantel, N; McPherson, K; Peto, J, and Smith, P G, 'Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples', *British Journal of Cancer*, 35(1) (1977), 1–39.

Pincock, S, 'Paul Meier (obituary)', *The Lancet*, 378(9795) (2011), 978.

Pohar, M, and Stare, J, 'Relative survival analysis in R', *Computer Methods and Programs Biomedicine*, 81(3) (2006), 272–278.

Rosner, G L, 'Bayesian monitoring of clinical trials with failure-time endpoints', *Biometrics*, 61(1) (2005), 239–245.

Rothman, K J, 'Estimation of confidence limits for the cumulative probability of survival in life table analysis', *Journal of Chronic Diseases*, 31(8) (1978), 557–560.

Ryan, T R, and Woodall, W H, 'The most-cited statistical papers', *Journal of Applied Statistics*, 32(5) (2005), 461–474.

Ryan, T R, and Woodall, W H, 'The most-cited statistical papers, (Most cited list in an Excel file (5/8/2011))', in M R Schell and J H Lee (eds), *Most cited papers*, Best Statistical Practice, 5 August 2011, http://beststatisticalpractice.com/most-cited-papers/ (accessed 25 March 2018).

Schultz, S, and Eisenberg, S, 'Statistician John W. Tukey dies', *News from Princeton University*, 26 July 2000, https://www.princeton.edu/pr/news/00/q3/0727-tukey.htm (accessed on 4 March 2018).

Smythe, B, 'Edward L. Kaplan (obituary)', *Newsletter of the Institute for Mathematical Statistics*, 35(10) (2006), 7

Stel, V S; Dekker, F W; Tripepi, G; Zoccali, C, and Jager, K J, 'Survival analysis I: the Kaplan–Meier method', *Nephron Clinical Practice*, 119 (2011), c83–c88.

Thompson, J D; Higgins, D G, and Gibson, T J, 'CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice', *Nucleic Acids Research*, 22 (1994), 4673–4680.

Tobacman, J, 'Paul Meier, P. 1924–2011. U. of C. professor applied statistics to medical research', *Chicago Tribune*, 18 August 2011, http://articles.chicagotribune.com/2011-08-18/news/ct-met-meier-obit-20110818_1_clinical-trials-research-experimental-treatment (accessed on 4 March 2018).

US National Archives, Personnel file of Kaplan, Edward L, Department of the Navy, The National Archives at St Louis, National Personnel Records Center, St Louis MO, USA (in accordance with the Freedom of Information Act (FOIA), USC 552(b)(6), copy dated 11 January 2016).

Van Noorden, R; Maher, B, and Nuzzo, R, 'The top 100 papers', *Nature*, 514(7524) (2014), 550–553.

Willett, J B, and Singer, J D, 'Applications of survival analysis to aging research', *Experimental Aging Research*, 17(4) (1991), 243–250.