NBER WORKING PAPER SERIES

DEATH OF THE SALESMAN, BUT NOT THE SALES FORCE: REPUTATIONAL ENTREPRENEURSHIP AND THE VALUATION OF SCIENTIFIC ACHIEVEMENT

Pierre Azoulay
J. Michael Wahlen
Ezra W. Zuckerman Sivan

Working Paper 24591
http://www.nber.org/papers/w24591

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2018

Address all correspondence to jmwahlen@mit.edu. Azoulay acknowledges the financial support of the National Science Foundation through its SciSIP Program (Award SBE-1460344). We thank seminar audiences at MIT and the University of Pennsylvania for useful discussions. The usual disclaimer applies. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Death of the Salesman, but not the Sales Force: Reputational Entrepreneurship and the Valuation of Scientific Achievement

Pierre Azoulay, J. Michael Wahlen, and Ezra W. Zuckerman Sivan

NBER Working Paper No. 24591
May 2018
JEL No. I23,O31

ABSTRACT

Using citations as a measure of valuation and death as a shock that affects efforts to "sell" scientific work but not the quality of the work itself, we estimate the importance of "reputational entrepreneurship" on the valuation of life scientists' research. Insofar as reputational entrepreneurship is impactful, it is unclear whether the most effective reputational entrepreneurs are those selling their own work ("salesman") or those promoting the work of others (the "sales force"). While the salesman has more incentive to promote her work, the sales force is larger and may be seen as more credible. We find that by commemorating the death of a scientist, the sales force boosts the valuation of the deceased's work relative to what the salesman could have done had she remained alive. This suggests that while science seeks to divorce the researcher's identity from their work, scientists' identities nonetheless play an important role in determining scientific valuations.

Pierre Azoulay
MIT Sloan School of Management
100 Main Street, E62-487
Cambridge, MA 02142
and NBER
pazoulay@mit.edu

J. Michael Wahlen
MIT Sloan School of Management
100 Main Street, E62-365
Cambridge, MA 02142
jmwahlen@mit.edu

Ezra W. Zuckerman Sivan
Massachusetts Institute of Technology
ewzucker@mit.edu

You have no control:
Who lives
Who dies
Who tells your story?

Lin-Manuel Miranda
Hamilton: An American Musical (2015)

1 Introduction

Insofar as we human beings have an impact on the world before we leave it, this legacy is effected via two possible channels: we manipulate the physical world, leaving behind artifacts or "products" that can be attributed to our agency; or, through our interactions with others, we influence their beliefs, preferences, feelings, or skills in ways that they might later recognize. Most people — and certainly those who take on public "producer" roles such as artists, politicians, or scientists — prefer that their legacy be significant and positive. But our ability to shape that legacy is limited. Once a product has left the hands of its producer, the objective features of the product are set. And once we are dead we can no longer promote our products or influence others' impressions of our impact on the world. Consider Winston Churchill's famous claim that "History will be kind to me for I intend to write it." Churchill undoubtedly had more ability to shape his legacy than most; but even during his lifetime his ability to shape his legacy was limited by his critics and opponents; and since he was mortal like the rest of us, he obviously did not get to have the last word.

In the same vein, sociological research on "reputational entrepreneurship" has documented how a politician's or an artist's legacy is affected by the producer's death, either by preventing the producer from playing the role of "salesman" in publicizing and promoting himself and his products or by influencing how other parties play the role of a "sales force" in publicizing and promoting their work (Bromberg and Fine 2002: 1139). In some cases, the death of the producer/salesman appears to have a negative effect on his legacy. For example, in accounting for why U.S. President Warren Harding is the "worst president of all time" (Holmes and Elder 1989), Fine (1996) notes that Harding was a reasonably popular president during his lifetime, and he was able to boast a strong economic and civil rights record; however, his early death in 1923 prevented him from defending his reputation in the wake of the Teapot Dome scandal while his erstwhile supporters had every incentive to let him take the blame. Yet while the death of the producer can have a negative impact on his

legacy, it can paradoxically have a positive effect insofar as it mobilizes a sales force composed of people who were positively influenced by the producer during her lifetime. Thus, Lang and Lang (1988) document how the sudden death of young etchers mobilized friends and family to commemorate the œuvre of the deceased, thereby making it less likely that the artist would be forgotten by the next generation. Fine (1996: 1188) too contrasts Harding's death with John F. Kennedy's, showing that Kennedy's supporters commemorated his life and work to such an extent that he became one of America's most popular presidents after his death, despite a rather brief and controversial term as president.

But to what extent does reputational entrepreneurship determine a producer's legacy? Whereas past research has focused on politics and art (e.g., Fine 1996; Jansen 2007; Lang and Lang 1988; Bromberg and Fine 2002; Kahl et al. 2010; McCormick 2015), the present study analyses how a scientist's death affects the amount of positive publicity that his papers receive, in the form of citations. Note well: the objective quality of a scientific paper obviously does not change when the scientist who authored the paper dies. Thus, insofar as the death of a scientist occasions a change in the citations to her papers, it follows that there is a significant subjective element in the valuation of science. In particular, if death leads to a reduction in citations, this would imply that scientists' "salesmanship" (broadly construed) plays a significant role in shaping its reception. And if death leads to an increase in citations, this would imply that the promotional efforts of the "sales force" of people that a scientist influences and inspires are key. Finally, to the extent that the effect of death on citations is lasting rather than fleeting, we gain insight into the extent to which the overall direction of science is significantly shaped by contingent, social factors.

To elaborate, there are two main reasons why science provides an especially good setting for advancing our understanding of how reputational entrepreneurship shapes social valuation generally and producers' legacies in particular. First, whereas in the artistic and political domains the producer's identity is considered relevant for evaluating their work, the "norm of universalism" that governs science (Merton 1979) implies that a scientist's reputation is irrelevant for evaluating her work, especially once it has passed peer review. Consider a key contrast between science and art: Whereas even a great a piece of art is not considered authentic unless its provenance is clear, it is the methodology of a scientific work that is thought to matter for scientific progress. Accordingly, Whitehead's dictum, "A science that hesitates to forget its founders is lost" is not applied to art. Relatedly, artistic works are generally presented and evaluated in the context of the artist's larger œuvre, and this body

of work is expected to have a consistent style (Sgourev and Althuizen 2014; Wohl 2017). Politicians too face significant penalties insofar as they are inconsistent (Barker and Carman 2010; Hummell 2010), and the reputations they build influence how we evaluate their actions. Thus, the observation that "only Nixon could go to China." Note finally that when death strikes a working artist or politician, it is unsurprising that this will affect the valuation of their work. The reason is that as long as a producer's works are evaluated in terms of their whole agenda, the meaning of this agenda remains open as long as the producer is still active.

By contrast, while the interpretation of earlier scientific papers can change in light of findings in subsequent research, it is irrelevant which specific scientist produced such research. In short, insofar as reputations should not matter for the evaluation of scientific papers, science represents an especially conservative setting for establishing the importance of reputational entrepreneurship. Of course, while it may be possible to ignore the reputation of the scientist when allocating credit to scientific papers, it is impossible and indeed undesirable when allocating *jobs* to scientists. As such, to the extent that we find that reputational entrepreneurship shapes the assessment of scientific papers, this would imply that science is more like art and politics than the norms of science would imply.

A second advantage of the setting of science is that it affords easier identification of the effects of reputational entrepreneurship. A key challenge in verifying any causal claim is to measure the impact relative to a counterfactual situation in which the event had not occurred (Lewis 1974). In politics, this is especially difficult because the number of observations is quite small and events are historically and contextually dependent. And identifying counterfactuals in art is challenging due to the absence of established criteria for judging pieces of art to be equivalent. In science, however, over 2.5 million articles are published annually after having completed peer review based on relatively consensual evaluation guidelines. As a result, it is possible to synthesize a counterfactual world in which death and/or reputational entrepreneurship did not occur by comparing articles with similar characteristics. In determining which characteristics are relevant, science's use of citations provides a further boon as citations represent an institutionalized means of recording and quantifying the community's (cumulative) assessment of a paper's quality over time (Merton 1988: 621). Thus, by contrasting citation trajectories, we can arrive at a precise estimate of the difference in the short and long-term assessment given to any two papers. Finally, memorialization events in science are fairly standardized, allowing the analyst to go beyond simply inferring the effect

of reputational entrepreneurship from death and allowing for disentangling the efforts of the salesman and the sales force.

To preview our findings, our analysis of elite academic life scientists shows that a scientist's death tends to provide a boost to the citation trajectory of their papers, and it does so by mobilizing scholars seeking to memorialize the deceased, thereby promoting her work and reputation posthumously. As a result, these scholars' research is better remembered than those still living, or those who died at a much older age. We also find that these effects appear to be permanent; for up to ten years after their deaths, their work continues to be cited more than comparable work by scientists who had not yet died. More specifically, we find that this effect is strongest for younger scientists who die suddenly; with respect to their least-cited papers at time of death. Overall, the results suggest that it is the sales force, rather than the salesman, which has more influence, and that the effect manifests itself in a shift in attention rather than valuation. More generally, the implication is that reputational entrepreneurship via the memorialization process affects valuations even in a context governed by the norm of disinterestedness and relatively clear quality criteria. Before reviewing our data, methods, and results, we clarify the theoretical issues at stake.

2 Theory

In this section, we look to clarify the conditions under which we might expect a scientist's death to have an impact on the trajectory of their citations. We begin by first examining the idealized conditions under which the reception of scientific work would be an unbiased measure of quality. We then relax those assumptions to clarify how reputational entrepreneurship could come to affect perceptions of quality, and the various ways this might manifest itself.

2.1 The "Ideal" Valuation Method for Scientific Works

It may at first seem unproductive to consider a naïve baseline by which scientific citations are an unbiased measure of quality. After all, few scientists are likely to assert that the community is completely objective in its valuation of research, especially as reflected in the number of citations to an academic paper. And yet one reason to consider this baseline is the great demand for objective measures of research quality — for the tenure process (Segalla 2008), for university and national rankings (Altbach 2012; Collyer 2013), and for

the awarding of prizes — citations are widely used because there is no consensus as to how they may be biased and no apparent alternative to using them. Moreover, by examining the idealized conditions under which citations are unbiased quality measures and then relaxing them, we gain a clearer view of when and why reputational entrepreneurship might be successful within science.

One idealized world in which citations would be unbiased measures of quality is a world governed by the Mertonian norms of universalism, communalism, organized skepticism, and disinterestedness (Merton 1979). In such a(n ideal) world, each scientific work would be objectively evaluated by the community at the time of publication (Merton 1968). Post-publication, the scientific community would know the objective quality of each work and would then build on the most promising articles, citing them to give credit where credit is due (Merton 1988). Thus, the best papers would receive the most citations, thereby making citation counts an accurate measure of quality.

Yet, this world has always been an unachievable ideal. Merton himself put the norms forward as science's aspirations, and not as descriptions of the actual practice of science (see Merton 1979). Additionally, a significant body of subsequent literature has shown that science often fails to live up to these ideals; for instance, recent work has found evidence of the Matthew Effect (e.g., Azoulay et al. 2014, Simcoe and Waguespack 2011), that scientists often choose research trajectories based on career interests as opposed to research interests (e.g., Foster et al. 2015), and more specifically, that citations are often used for a variety of reasons beyond simply giving credit (e.g., Leydesdorff 1988).

A second, and more plausible, basis for expecting citations to be unbiased measures of quality derives from an appreciation for the factors that make science a relatively "efficient market." To clarify, let us consider the analogy of the stock market, and what is responsible for deviations from efficiency (Brav and Heaton 2002; Zuckerman 2012b). Under the efficient market hypothesis, stock prices are said to be unbiased indicators of value that reflect all publicly available information about a company (Fama 1965, 1970; Malkiel 1989). There are two key mechanisms underlying this idea, the first of which is arbitrage — that is, that insofar as there is a gap between the "fundamental" or "intrinsic" value and the current price, investors can profit from the difference by buying low and selling high or (short-)selling high and (re-)buying low (Zuckerman 2012b: 227-228). The second key mechanism is learning: as some investors make great profits from arbitraging price and value, other investors either

suffer from capital erosion or they learn the methods employed by the former (Zuckerman 2012b: 228). The efficient markets hypothesis holds that the profits — and thus the incentives — are so great that this process works extraordinarily quickly and expeditiously such that at any one time, prices are unbiased estimates of future cash flows — i.e., of fundamental values.

Applied to the context of science (see Zuckerman 2012a: 237-239), the idea that citations are an unbiased estimate of scientific quality hinges on the assumption that the arbitrage and learning mechanisms operate strongly. An example of scientific arbitrage is when a scientist identifies an undervalued idea and builds on it, or when she sees a paper valued too highly and seeks to attack it and thereby pushes her field in a different direction. In these scenarios, the scientist is presumed to be motivated not by fealty to communal norms but by self-interest. The incentives for developing undervalued ideas are strong to the extent that scientific careers are made by finding and/or recombining ideas that were overlooked (Uzzi et al. 2013). And the learning process is clearly important as well; science is continually beset by the steady advance of new ideas and the falsification of old ones (Popper 1959), and (though less often) paradigm shifts which lead scientific fields in completely new directions (Kuhn 1970).

Thus, to the extent that either the Mertonian norms govern science and/or the arbitrage and learning mechanisms in science are strong, a scientist's death should have no impact on the citation trajectory of the papers they authored before they passed away. This leads us to formalize a proposition, which we consider a naïve but important baseline:

> **Proposition 0:** A scientist's death will have no effect on the citation trajectory of their work.

2.2 Valuation Entrepreneurship

But, just as arbitrage (and therefore learning) is limited in the stock market (Zuckerman 2012b), so it is in science as well (Zuckerman 2012a: 237). In general, the most effective forms of arbitrage are what Zuckerman (2012a) calls "valuation opportunism." This occurs when an actor who believes an asset to be mispriced can profit from that discrepancy without any regard for others' opinions. In the stock market, this is the canonical case of a value investor such as Warren Buffett who buys a public company outright and earns income from

it at a discounted price (if they are right that the company was undervalued) (see Graham 2006 [1984]). In science, however, this strategy is much more limited. While some ideas (such as penicillin or TNT) have a clear immediate use such that those who adopt them early can benefit from them even in the face of widespread skepticism, the value of most ideas is hard to assess and apply without significant growth in the community of supporters both inside and outside academia.

Similarly, while financial markets can learn from the removal of misinformed investors quickly (e.g., through margin calls), scientific progress takes much longer. This is a result of a number of factors, for instance, the nature of the scientific method, the review process, and the fact that established scientists with a vested interest in a particular paradigm are likely to prevent others from attacking or modifying it, giving rise to Planck's famous dictum that science only advances "one funeral at a time" (Azoulay et al. 2016). Consequently, a scientist who discovers that a well-respected idea is overvalued faces a dilemma in that while revealing its flaws may well be the best long-term strategy, it may take years to bear fruit (Foster et al. 2015). This likely creates a general reluctance to try out undervalued ideas and research modalities as scientists fear their efforts will not be rewarded in time for their careers to progress (Foster et al. 2015).

It is in the context of the limits to arbitrage and learning that valuation entrepreneurship becomes a salient, and sometimes necessary, strategy for "market" participants.[1] Were scientists able to use valuation opportunism to earn timely rewards as the field re-appraised work in an unbiased manner, scientists would have little need to promote their work. Similarly, value investors prefer to remain silent about their investments to avoid others exploiting them and thereby reducing their returns (Zuckerman 2012a: 235). Absent a mechanism such as this one, however, scientists are forced to act as valuation entrepreneurs who must change other participant's standard of quality in order to reap the rewards of their investment. For this reason, much like short-sellers, scientists who disagree with the field's current valuations must look to promote work they agree with (known as "talking their book" in the parlance of finance) in order to bring other participants to their point of view, thereby causing the field to reflect their own valuation (Zuckerman 2012a: 235; Botelho 2017).

[1] As Zuckerman (2012a:235) notes, "valuation entrepreneurship" differs from "valuation opportunism" in that it connotes contrarian strategies which explicitly rely on the market changing its valuation of a given security, while "valuation opportunism" allow the investor to profit regardless of whether or not the market itself changes its valuation, as in the example of Warren Buffet above.

The manner in which valuation entrepreneurship shifts valuations is likely to differ between financial markets and science, however. Financial markets have relatively low search costs, as participants are largely aware of the options before them and have likely formed some opinion of them. In such a context, valuation entrepreneurship is most likely to occur when one participant successfully changes the conclusions another participant has drawn regarding a specific asset. In science (much like art), however, the search costs are much higher as the number of works produced is greater than the participants can be aware of at any given moment (Cole 1970; Evans 2010). Consequently, valuation entrepreneurship can occur not only when a participant's conclusions are changed about a specific idea, as in financial markets above, but also when participants become aware of work unfamiliar to them. Raising awareness of a work effectively raises its value as works with the attention of the field are eligible for either a high or low valuation, while those with little to no attention can only be forgotten, effectively being valued at the bottom of the field (Denrell and Le Mens 2016). This occurs most dramatically in art, as works by well-known artists are more highly valued than technically superior works by completely unknown artists (Lang and Lang 1988). Yet, this also occurs in science as awareness of an idea is a pre-condition for credit, which Merton describes as the "coin of the realm" (Merton 1968). Thus, scientists are much more likely to build off a popular line of work even if there are other works which are more promising by virtue of the fact that they unable to know every paper's value at any given time (Boudreau et al. 2016; Iaria and Waldinger 2015; Mulkay 1972).

2.3 Salesman versus Sales Force

But how would scientists most effectively influence the community's perception of work? Perhaps the most obvious manner is through aggressive self-promotion. Scientists who believe their research is undervalued by the community may seek to raise awareness of it through press releases, teaching graduate courses, presenting at conferences, etc. This implies that at any given point in time, the level of citations a paper receives is a function of the quality of the paper and the amount of salesmanship it has received. Thus, after the death of the scientist, the latter factor is removed, and therefore, the number of citations should decline. This results in our first proposition:

> **Proposition 1:** Insofar as the reception of scientific work increases through efforts of self-promotion, the death of a scientist will cause the reception of scientific work to decrease relative to the work of extant scientists.

Yet, the efficacy of self-promotion may be limited for two reasons: size and motive. First, and most obviously, any given scientist's capacity for promoting her work is dwarfed by the number of colleagues who could promote it. Second, because the scientific community is aware that authors have a strong incentive to see their work more highly valued, efforts to increase one's own idea's valuation may be overshadowed by concerns of partiality. Thus, while the salesman may have the most incentive to sell their work, their credibility may be so low as to make them ineffective.

In such a context, touting the ideas of others may seem more credible than touting one's own ideas. When there is not an obvious career benefit to promoting a given idea, the field may perceive such promotion as stemming from the promoter's assessment of the work's intrinsic quality and therefore to be more receptive to the message. This raises the question of when one would be most likely to find a large group of individuals touting the work of another. Lang and Lang (1988) suggest that this type of effort is most likely to occur after a death, when supporters are brought together by the tragedy and seek to record the person's life and work as a tribute to the deceased. This leads to the creation of "memory events" — biographies, news articles, and exhibits of their life and œuvre (Lang and Lang 1988: 94). As a result of these efforts, the field's attention is directed towards the work of the deceased, thereby raising its valuation in the manner described above (Lang and Lang 1988: 97). In the field of etchers, Lang and Lang claim that this was effective to such an extent that memorialized etchers were remembered vastly beyond their living counterparts, even those with superior work (Lang and Lang 1988: 97). Consider their example of Elizabeth Fyfe, below (Lang and Lang 1988: 93):

> One other case in point: Elizabeth Fyfe, who died in Switzerland in 1933, just after her thirty-fourth birthday after a long bout with tuberculosis, had been hailed by British critics as "one of the most original and accomplished young etchers." That her name and her work, which amounted to just over 1,600 impressions, somehow survive, whereas those of others once equally or better known do not, has much to do with her premature death. Her teachers, her friends, her collectors, and other etchers rallied, while she was in the hospital, to organize an exhibition of her work, complete with catalog, and then used the proceeds from sales to help pay for the care she needed. Her dealer saw to it that her plates were printed when she could no longer do so herself and gave a full set of her prints to Fyfe's sister. In this way, the many persons mobilized by the tragedy helped to preserve the work and, thereby, to sustain the memory of the artist.

We label this mechanism (the creation of memories about a deceased author) "memorialization" and the supporters that participate (or decline to participate) in this process the "sales force." This results in our second proposition:

> **Proposition 2:** Insofar as the reception of scientific work increases from the efforts of supporters and the death of a scientist increases the rate of memorialization by supporters, the death of a scientist will increase the reception of his or her work relative to extant scientists.

2.4 Predictions Distilled

In summary, there are three different theoretical predictions as to what might happen to the perception of a scientist's work after their death, depicted graphically in Figure 1, below. In these stylized graphs, a flat gray line runs the length of each representing the number of citations a given deceased scientist's publication would have received had they remained alive. Under the assumptions of the ECH, this line also represents the predicted effect of death on citations. In panel A, the prediction given by the importance of the salesman is depicted. According to Proposition 1, the trend of citations (represented by a dashed blue line) decreases after the death of the scientist, as she is no longer alive to sell her work. In panel B, the dashed blue line slopes upward, representing the boost in citations the scientist receives due to the mobilization of the sales force. Determining which of these theoretical predictions best represents the data is the central empirical question of this paper.

As suggested in Panel C, however, if the sales force fails to mobilize it creates an indeterminant outcome. Under this scenario, the result would be no change in citations after death which is empirically indistinguishable from the ECH and from the effects of the salesman and sales force counteracting each other. To address this possibility, we examine the process of memorialization, represented by the red box in Panel B of Figure 1. By collecting the documentation of scientists' deaths created by their followers, we look to measure the extent to which the scientists were memorialized and to estimate the relationship between memorialization, the circumstances of the death, and the posthumous reception of their work. In this manner we are able to develop a robust measure of the efficacy of both the sales force and the salesman and discern if the two cancel each other out or if neither has any effect at all.

Insert Figure 1 Here

3 Data and Empirical Design

The design of our empirical analysis unfolds in three separate steps. The first step is a *causal* analysis: how does the premature death of an eminent biomedical academic researcher change the recognition (as measured by citations) of her work, compared to the work of other eminent researchers who do not die prematurely. The level of analysis for this step is an article/scientist pair, and the main challenge to be overcome is the building of a control group of articles that plausibly pin down the citation trajectories of the deceased scientists' articles had they remained alive. The second step is *descriptive*: what are the correlates of memorialization? In this step, the level of analysis is the individual scientist and the key challenge to be overcome is the measurement of the memorialization process, which is highly variegated and would, at first blush, appear to defy efforts at quantitative reduction. The third and final step ties the causal and descriptive analyses together. We ask whether the memorialization process is a plausible mechanism through which scientific work gets remembered in the long run. The main challenge is one of *prediction*: for each article, we must be able to forecast the citation trajectory that would have been observed if the scientist had remained alive, so as to isolate a net citation premium (or deficit) for each deceased scientist. With these forecasts in hand, we can then examine whether variation in memorialization intensity correlates with extinction-induced "excess" citation rates.

Below, we provide a detailed description of the process through which we assembled the data set used in the statistical analysis. We begin by describing the criteria used to select the sample of elite academics, with a particular focus on the timing and the manner of their deaths. The focus then shifts to the publications deceased and still-living scientists authored during their lifetime, and how one might build a matched sample of publication/scientist pairs where the citations received by articles authored by extant scientists offer a plausible counterfactual to the citations that articles authored by extinct scientists would have received had not died prematurely. Finally, we document how we measured the memorialization process for each individual scientist. Throughout this description of the data, we outline how the construction of the sample addresses the empirical design challenges enumerated above, while leaving the details of our statistical procedure to section 5.4.

3.1 Institutional Context

Our empirical setting is the academic life sciences. We focus on this domain for three reasons. First, its sheer size. U.S. Medical Schools employ over 150,000 faculty members (twice the number of physical scientists) and this figure underestimates the size of the labor market since it does not take into account scientists and engineers working at NIH, in non-profit research organizations (such as the Salk Institute), for independent hospitals (such as the Cleveland Clinic), or within Schools of Arts and Sciences (such as MIT, UC Berkeley, or Rockefeller University). Academic biomedical research also garners over 70% of all non-defense Federal R&D dollars. The large size of the labor market is important for reasons of statistical power: our key source of variation is generated by the premature death of eminent scientists, and these events are relatively rare. Importantly, the members of this labor market share broadly similar norms, career goals, incentives, and operate within comparable institutional structures.

Second, scientific discoveries over the past half-century have greatly expanded the knowledge frontier in the life sciences, and these advances have resulted in more specialization, as well as an increase in the size of collaborative teams (Jones 2009; Wuchty et al. 2007). These trends help ensure that career shocks only affect relatively narrow swathes of the intellectual landscape. Were our research domain smaller in size, or less balkanized across narrow subfields, it would be challenging for us to identify control articles or control scientists (Azoulay et al. 2010).

Third, and perhaps more pragmatically, our setting is blessed by an abundance of data sources. The careers of eminent, still-living life scientists are extensively described in publicly-available curriculum vitas, *Who's Who* profiles, or laboratory web sites. Deceased scientists leave in the wake of their passing an extensive paper trail in the form of biographical articles, reminiscences authored by former colleagues, and obituaries. We combine these data with large-scale databases such as the Faculty Roster of the Association of American Medical Colleges (AAMC), the free and publicly-available bibliographic database *PubMed*, NIH's Compound Grant Applicant File (CGAF), and citation information from the *Web of Science*. Together, these sources of information allow us to create an accurate longitudinal record of publications, citations, and funding for each scientist in the sample.

3.2 Sample of Elite Academic Life Scientists

In this context, our empirical approach uses the death of elite scientist as a lever to estimate the extent to which the reception of their work changes posthumously. Our focus on the scientific elite is justified pragmatically and substantively. Pragmatically, elite scientists leave behind a large body of work as well as colleagues with an interest in the preservation of their legacy. This makes tracing and documenting their careers much more viable than with less-known and less-published scientists. Additionally, the distribution of publications, funding, and citations are extremely skewed (Lotka 1926; de Solla Price 1963) as only a tiny fraction of scientists contribute to the advancement of science (Cole and Cole 1972; Zuckerman 1967).

Rather than a limitation, we believe that our focus on the scientific elite is substantively justified in light of our goals. One would expect the articles of eminent scientists to be identified and evaluated immediately after their publication, relative to the articles authored by scientists of lesser repute. This should in turn should make it less likely that reputational entrepreneurship should matter here. To some extent, this is testable since our metrics of eminence are chosen such that substantial heterogeneity will exist even within the sample of eminent scientists.

We began by demarcating a set of 12,935 "elite" life scientists (roughly 5% of the entire relevant labor market) who are so classified if they satisfy at least one of the following criteria for cumulative scientific achievement: (a) highly funded scientists; (b) highly cited scientists; (c) top patenter; or (d) member of the National Academy of Sciences. Because these four measures rely on achievements over the course of a scientist's career, they will tend to select older scientists. To create more demographic balance, we add three additional measures that capture individuals with promise at the early and middle stages of their scientific careers (regardless of whether that success endures): (e) NIH MERIT awardees; (f) Howard Hughes Medical Investigators; and (g) early career prize winners. Appendix A provides additional details regarding these seven metrics of "superstardom."

We trace back these scientists' careers from the time they obtained their first position as independent investigators (typically after a postdoctoral fellowship) until 2006. We do so through a combination of curriculum vitae, NIH biosketches, Who's Who profiles, accolades/obituaries in medical journals, National Academy of Sciences biographical memoirs, and Google searches. For each one of these individuals, we record employment history, degree

held, date of degree, gender, department affiliations, as well as complete list of publications, patents, and NIH funding obtained in each year.[2]

The next step in the sample construction process is to subset from this pool scientists whose premature death will "treat" their past output, as well as those scientists who could potentially serve as controls. First, we select scientists whose death intervenes between 1969 and 2003 (for the treatment group), as well as those who were still alive in 2006 (the end of the observation period).[3] Second, we need to ensure that the treated scientists had not entered a pre-retirement phase of their career. This is trickier, because the timing of retirement is endogenous, and scientists who do not wish to retire can show great initiative in subverting the rules surrounding mandatory retirement. To overcome this challenge, we make full use of the narrative data contained in the dossiers we compiled for each scientist (deceased or not); we also examine publication output as well as funding received to weed out from the sample those who either "meaningfully" retired or whose output shows sign of abating prior to their death or the end of the observation period.[4]

As a result of these steps, we identify 676 "treated" scientists (see Table 1). The mean age at death is 63, with the youngest scientist dying at age 35 and the oldest dying at age 89.[5] We then investigate the cause of death in this sample of 676 scientists to classify their deaths as being either "sudden" or "anticipated." This is less difficult than it appears, since most obituaries typically are quite specific in this respect.[6] To distinguish sudden from anticipated deaths, we use an arbitrary distinction between deaths that likely occurred with six months notice or less, versus those that likely occurred with more than six months notice. In practice, this sudden category mostly comprises fatalities due to heart attacks, car accidents, and sudden onset illnesses. Conversely, most "anticipated" deaths are from various

[2]Appendix B details the steps taken to ensure that the list of publications is complete and accurate, even in the case of stars with frequent last names.

[3]The control scientists might well die or retire, but only after 2006. An implication of this design choice is that even for the scientists who die "late" (e.g., in 2003), we will have at least three years of citation data to pin down how their passing changes the recognition of their work.

[4]In previous work, one of us has verified that it is essentially impossible to predict death in this sample using measures of lagged publication output (Azoulay et al. 2010).

[5]How can one die at the age of 89, and one's passing still be deemed "premature?" Easily, as it turns out. Audrey Gorbman (1914-2003), described in academic obituaries as the "father" of the field of comparative endocrinology (Bern and Sower 2003), succumbed to Parkinson's disease but still published two first-authored article in the last year of his life.

[6]We exclude from the sample one scientist who took his own life, and a further two for whom suicide was hinted at. In some instances, where the cause of death could not be ascertained from the obituaries, we contacted former collaborators individually to clarify the circumstances of the superstar's passing. We were unable to ascertain the cause of death for 29 (4%) of the 676 treated scientists.

forms of cancer, or other long-term illnesses. In the treated scientist sample, 307 (45%) of the treated scientists died suddenly death, while 340 (50%) died from an anticipated illness.

Table 1 provides descriptive statistics on the sample of deceased eminent academics (see Appendix E for a complete list of these scientists, along with basic demographic and achievement data). The overwhelming majority (91%) are males (patterns of entry by gender into the underlying labor market have only recently equalized, and our sample reflects the extreme gender imbalance that prevailed for most of the time period we study). Of note is the fact that even within this sample, substantial variation in recognition exists: whether one measures eminence through publications, NIH funding, or citations (excluding those citations that accrue after the scientist has passed), the mean is always much higher than the median.

Insert Table 1 Here

3.3 Matched Sample of Articles

The scientist level of analysis is not well-suited to the challenge of identifying the causal effect of death on the reception of an academic's work. To be sure, we could try to construct a synthetic cohort of live scientists who look otherwise similar at some point of time to the scientists who die prematurely; but, any such comparison would conflate the effect of death with the effect of being able to build on one's previous work in a cumulative fashion. In contrast, the article-level of analysis, as well as the flow of citations that accrue to each individual article over time, does provide a useful source of variation. This is because the content of a given article is fixed over time, whereas the engagement of the audience with it can (and does) change over time. Moreover, there is a very natural datum that determines unambiguously a "before" and an "after" period for each article: the timing of its author's death.[7]

However, a simple difference between citations that accrue to a paper after, rather than before, the death of a scientist is not enough to yield estimates with a plausibly causal

[7]This basic insight is not new. For instance, Murray and Stern (2007) ask how citations to articles shift once the underlying results appear in a patent; Azoulay, Stuart and Wang (2014) ask how the receipt of an accolade changes the citation trajectories of articles that appeared before the accolade was received; Azoulay, Graff, Zivin, and Sampat (2012) investigate how the mix of local to non-local citations changes after a scientist moves to a geographically distant institution.

interpretation of the effect of a scientist's passing. This is because the memory of any article (or scientist) must eventually fade. Examine (in Figure 2) the mean number of annual citations received by the 676 treated scientists in the sample, both before and after the death. This peaks the year before the death, and then undergoes a decline that is steep, though it will take close to 40 years for the memory of any work by a deceased scientist to disappear from the scientific literature. Therefore, the question is not whether the recognition given to the work of deceased scientists will decrease after they die, as it surely will. Rather, the challenge is to assess this decline relative to the citation trajectory of papers in general, and more precisely relative to papers whose recognition potential was similar at the time of the scientist's passing. Therefore, we need to construct a group of control articles that can plausibly capture this counterfactual.

Insert Figure 2 Here

Our approach is to recruit control articles from the vast set of articles authored by elite scientists who did not die prematurely. There exist uncountable ways to build a control sample in this way, and this step necessarily entails some degree of judgment. For this reason, it is valuable to specify the characteristics that one would want the control articles to exhibit if at all possible. First, one would them to be published contemporaneously with the "treated" articles; Second, one would like them to be of similar expected impact and fruitfulness, relative to the treated article, up to the time of death; Third, the scientist who published the control article should be similar, from a demographic standpoint, to the treated scientist; And fourth, the control article should be isolated (in "intellectual space," but maybe also in "social space") from the treated scientist. In practice, it is impossible to identify for each treated article a "fraternal twin" that matches it exactly on a list of author and article characteristics.

Pragmatically, we specify a handful of covariates along which matched article/cite pairs must resemble each other, and we implement a coarsened exact matching procedure (Iacus et al. 2011) to identify all the articles among those published by live scientists that satisfied these criteria (so that each treated article can and typically does have more than one associated control article). This is best explained through an example, but we must first address an obvious obstacle. Modern science is a team sport, with rates of coauthorship that have steadily increased over the past 40 years (Wuchty et al. 2007). With long authorship rosters (the median number of authors in the PubMed universe was 4 in 2002), how can

we know the author with which the article is most closely associated? Here, we are helped by a strong norm in biomedical research which invariably puts the principal investigator on a research project in last authorship position on any paper that results from the funding s/he was able to mobilize (Gans and Murray 2011). Therefore, below we will focus only on original research articles where the focal scientist appears last on the authorship roster.[8]

3.3.1 Example of a Treated-Control Pair

Consider the paper "NMDA Receptor Losses in Putamen from Patients with Huntington's Disease," published in the journal *Science* in 1988 originating from the laboratory of John B. Penney, Jr., an eminent Harvard neurologist who died in 1999 from heart failure. Using the coarsened exact matching procedure described in detail in Appendix C, we can match 15 publications to this article, also published in *Science* in 1988, and where a superstar who did not die was in last authorship position. Figure 4 illustrates the matching with one of these articles, "Identification of a Putative Regulator of Early T Cell Activation Genes," which came out the laboratory of Gerald Crabtree, a pathologist at Stanford. By 1999, the Crabtree paper had garnered 410 citations, which is twice as many as the Penney paper had received. Yet, both articles belong to the top percentile of the 1999 citation distribution for the universe of papers published in 1988. Notice as well that Crabtree and Penny were born in the same year. This is not happenstance, as the matching procedure selects for articles whose lead author is at most three years younger, or older, than the lead author of the treated article. The number of authors for each of the paper is identical, and a close match in terms of authorship roster length is also one of the criteria we use.

Yet, it is clear that there are still observable differences between these articles. The two lead authors do not match particularly closely on metrics of achievement, for example. Nothing would prevent us from extending the list of match covariates to include measures such as cumulative funding or citations. But this would entail being unable to find a match for many of the treated articles. As a result, we have focused on a small set of covariates at the article-level (journal, year of publication, and citations received up to the year of death) and only one scientist-level covariate that struck us as important given the context: the age

[8]To be sure, we can have a deep imprint on a research project and yet occupy authorship position other than last. In the case of inter-lab collaboration, for instance, it is not unusual to observe one of the PI occupy the first authorship position, or the next-to-last position. What is important for our purposes is that it is difficult to imagine circumstances where an author does occupy the last author position and s/he is not closely identified with the work.

of the scientist. One might have chosen to match on a host of other scientist-level covariates: eminence at death (for example, measured by citations), gender, or degree (MD vs. PhD). From an empirical design perspective, the lack of balance is only a threat to identification when one has a reason to believe that the unbalanced covariate is correlated with treatment (the timing of death, in this study). We will come back to this issue when presenting the results.

Two additional about this pair of articles are worth mentioning, since they hold true more generally in the sample. Crabtree and Penney never collaborated while Penney was alive. Furthermore, these two papers tackle unrelated topics. Formally, PubMed, our data source for biomedical publications, does not list one as being topically related to the other (The PubMed Related Article Algorithm will be described in more detail below). This is important insofar as a desirable feature of the control group is to be unaffected by the treatment event. By eliminating articles by collaborators as well as topically-related articles from the list of eligible controls, we bolster the claim that the control articles can pin down a credible counterfactual citation trajectory.

Insert Figure 4 Here

3.3.2 Descriptive Statistics

The procedure described above yields a total of 128,591 papers authored by 6,782 control scientists, as well as 18,523 treated papers authored by the 676 deceased scientists. Table 2 provides descriptive statistics for control and treated publications in the baseline year, i.e., the year of death for the deceased scientist. A number of the covariates are balanced between treated and control publications solely by virtue of the coarsened exact matching procedure — for instance, the year the article was written, the number of authors, and the number of citations at the time of the (counterfactual) death. However, covariate balance in the level of eminence at the time of (actual or counterfactual) death for treated and control scientists (measured through NIH funding, number of articles published, or cumulative number of citations) was not guaranteed by the matching procedure.

Insert Table 2 Here

Figure 3 examines differences in the shape of the distribution for citations received by treated and control articles, respectively. Overall, they are quite similar, particularly in the upper tail. On the other hand, the distribution of control articles exhibits slightly more mass in the bottom quartile, relative to the distribution of treated articles. Of course, balance in the *stock* of citations at baseline is not technically required for the validity of the empirical exercise. More important is the absence of a trend in the *flow* of citations up until the time of treatment. An important step of the empirical analysis will be to verify, *ex post*, the absence of pre-treatment trends.

Insert Figure 3 Here

3.4 Memorialization Data

To study the process of memorialization, we focus on the sample of treated scientists only. We remove from it the 29 stars whose cause of death remains unknown (i.e., not confirmed to be either sudden or anticipated), resulting in a sample of 647 scientists who died prematurely. Through systematic web and PubMed searches, we collected and hand-coded all "memory events" for each of these scientists in our sample. These include obituaries and reminiscences that appeared in medical journals, obituaries that appeared in newspapers, festschrifts, symposia, National Academy of Science Biographies, Wikipedia pages, university press releases, and major awards (e.g., a field-wide award for best paper in a given field named after a deceased scientist). Appendix D provides further details, but the main descriptive statistics are displayed in Table 1. We code an average of four "memory" events for the scientists in this sample. 31% of these events take place in the year of death, 32% in the year after the death year, 9% two years after the death year, and only 3% three years after death. In other words, by three year after death, the flow of memorialization events reduces to a trickle.

Since our period of observation for the death of eminent scientists is between 1969 and 2003, one might worry that certain data sources become available only later in the sample, mechanically producing more memorialization events for scientists who died more recently. Our main results, however, pertain to identifying the correlates of *academic* memorialization events — those that appeared in scientific journals — and there is less reason to suspect that the emergence of the internet led to an increase in the number of these types of events.

To guard against this peril, we will include in all our regressions a full suite of indicator variables for the years of death in the data.

3.5 Predicting Long-Run Posthumous Recognition

The last step of our analysis examines empirically whether memorialization efforts are a plausible intermediate outcome in the relationship between a scientist's death and long-run acclaim. It would be intuitive to simply modify the citation analysis by substituting citation outcomes with measures of memorialization activity. This is impossible, however, since measures of memorialization are only available for prematurely deceased scientists. We follow an alternative approach to generate measures of citation premium or deficit (whichever the case maybe) for each treated scientist.

The computation of these measures begins with an acknowledgement that the difference-in-differences modeling strategy explained above, while well-suited to the challenge of establishing the causal effect of premature death on citation trajectories, is not adapted to the task of *predicting*, at the article level, the future time-path of citations.[9] To generate article-level predictions, we begin by collapsing the data in the longitudinal dimension, such that for each article (treated or control) there are exactly two observations, one before the year of death or counterfactual death, and one from the year of death onwards.

We then run a very simple negative binomial model where *post-death* citations for paper i authored by scientist j are regressed on the log of *pre-death* citations for article i, a treatment indicator which is equal to one if scientist j died, and a vector of control covariates X_{ij}:

$$E\left[cites_{ij}^{AFTER}|X_{ij}\right] = cites_{ij}^{BEFORE} \cdot exp\Big[\alpha_0 + \alpha_1 TREAT_j + \alpha_2 X_{ij} + \nu_i\Big]$$

where ν_i is an omitted variable such that e^{ν_i} follows a gamma distribution with mean 1 and variance ζ (Cameron and Trivedi 2013: pp. 80-89). The vector X includes a set of indicator variables for article i's number of authors, a set of indicator variables for article

[9] In fact, the precise statistical procedure we will use, the quasi-maximum likelihood fixed effects Poisson estimator due to Hausman, Hall, and Griliches (1984) only allows us to characterize how scientist death shifts the conditional mean of the flow of citation over time. It would be invalid to use the resulting estimates to compute a prediction for each article in the sample. Yet, it is the appropriate estimator for the causal analysis because it will generate consistent estimates under mild regularity assumptions (Wooldridge 1997). Other count data estimators, such as negative binomial or zero-inflated count models are much more well-suited to the task of prediction, but these predictions are only valid if their underlying distributional assumptions are correct.

i's year of publication, and a set of indicator variables for scientist j's year of death (or counterfactual death). The predicted counts from this simple regression are then computed to generate a measure of *expected* citation. Note that the model includes an offset for the number of citations that had accrued to article i up to and including in the year of death. In other words, our predicted citation rate accounts for any momentum that article i's citation trajectory portended at the time of death.

"Excess" citation is simply measured as the difference between the predicted count and the actual number of citations received by the article after the year of death.[10]In a final step, we sum both the measure of expected citation and excess citation across all articles written by each scientist to generate an individual measure of citation premium (or deficit). Figures 6a and 6b provide an histogram for the distribution of both measures. Note that the mean of the excess citation measure is 24, while the median is -1.

Insert Figures 6a and 6b Here

In the cross-section of extinct scientists, we then run simple OLS specifications where expected or excess citations are regressed on the intensity of memorialization activities for each scientist, as well as a handful of control variables (such as gender, highest degree, cause of death, age at death, and year of death indicator variables).

4 Results

The exposition of the statistical results proceeds in stages. After a brief review of methodological issues, we provide results that pertain to the main effect of a scientist's death on the reception of their work. Second, we attempt to elucidate the mechanism (or set of mechanisms) at work to explain our most robust finding, that papers by scientists which die young and suddenly see a large and sustained increase in citation rates relative to those of scientists who are still living. We do so by measuring the extent of memorialization of deceased scientists and using it to predict posthumous citation rates. Finally, we correlate predicted excess citations received with the intensity of memorialization activity at the individual level.

[10]This measure of excess citations can, and in fact often is, negative.

4.1 Statistical Considerations

Our estimating equation relates the effect of a scientist's death on citations in the following way:

$$E\left[cites_{it}|X_{it}\right] = exp\Big[\beta_0 + \beta_1 AFTER\_DEATH_{jt} + \beta_2 AFTER\_DEATH_{ijt} \times TREAT_{ij} + f(AGE_{it}) + \delta_t + \gamma_i\Big]$$

where $cites_{it}$ is the number of citations paper i receives in year t (purged of self-citations), $AFTER\_DEATH$ denotes an indicator variable that switches to one in the year after the superstar (real or placebo) associated with i passes away, $TREAT$ is an indicator variable set to one if the scientist dies during the period, $f(AGE_{it})$ corresponds to a set of indicator variables for the age of article i at time t (measured as the number of years since the year of publication), the δ_t's stand for a full set of calendar year indicator variables, and the γ_i's correspond to article fixed effects, consistent with our approach to analyze *changes* in the flow of citations within each article following the passing of an elite scientist.[11]

We follow Jaravel et al. (2018) in including in our specification an indicator for the timing of death that is common to treated and control articles (whose effect will be identified by the coefficient β_1) in addition to the effect of interest, an interaction between $AFTER\_DEATH$ and $TREAT$ (whose effect will be identified by the coefficient β_2). The effects of these two variables are separately identified because: (i) deaths are staggered across our observation period and (ii) control publications inherit a counterfactual date of death since they are uniquely associated with a treated publication through the matching procedure described in section 4.2. The inclusion of the common term addresses the concern that age and calendar year fixed effects may not fully account for shifts in citation activity around the time of the scientist's passing. If this is the case, $AFTER\_DEATH$ will capture the corresponding transitory dynamics, while $AFTER\_DEATH \times TREAT$ will isolate the causal effect of interest. Empirically, we find that in some specifications, the common term has substantial explanatory power, though its inclusion does not radically alter the magnitude of the treatment effect.

Estimation. The dependent variable of interest, citations accrued per year, is skewed and non-negative. Specifically, 51.56% of the articles receive no citations in a given year while

[11] To avoid confusion, we have suppressed any subscript for the scientist. This is without loss of generality, since each article is uniquely associated with a single scientist (i.e., there can only be one individual in last-authorship position for each article).

0.02% accumulate over one hundred. Following a long-standing tradition in the study of scientific and technical change, we present conditional quasi-maximum likelihood estimates based on the fixed-effect Poisson model developed by Hausman et al. (1984). Because the Poisson model is in the linear exponential family, the coefficient estimates remain consistent as long as the mean of the dependent variable is correctly specified (Gouriéroux et al. 1984). We cluster the standard errors at the scientist level in the results presented below.

Outcome variables. Our primary outcome variable is the rate of yearly citations (net of self-citations). In order to better understand the activities of the "sales force," however, characterizing the relationship between citing and cited articles is also of interest. Specifically, are posthumous citations more likely to come from former collaborators or trainees? Are they more likely to originate from within the narrow subfield of the cited article, or from outside that narrow subfield? Or are they more likely to be circumscribed in geographic space, for example emerging from authors employed by the same institution as that of the deceased scientist? We parse all the citing-to-cited article pairs to distinguish between such relationships in social space, intellectual space, and geographic space.[12] We then aggregate these data up to the article-year level to compute citation counts from related versus unrelated authors.

Subsamples. The set of articles whose citations we analyze comprises all "original" (i.e., excluding reviews, comments, editorials, etc.) articles published by elite scientists (treated or control). The importance of each of these articles varies widely in the *within-scientist* dimension of the data. We will therefore run the citation analysis on the overall sample, as well as on subsamples designed to highlight the effect of death on different parts of the distribution of scientific impact. To do so, we assign each original article the percentile of the citation distribution to which it belongs, given its vintage. When computing these empirical distributions, we also take into account the year of death or counterfactual death. This allows us to compare between the citation impact of each article in the sample, regardless of the year in which it appeared and regardless of the time of treatment.

[12] Appendix B describes how this is achieved. Briefly, matching each author on citing and cited articles with the Faculty Roster of the Association of American Medical Colleges (AAMC) allows us to distinguish between publications with and without former collaborators or trainees, and with or without authors colocated with the focal elite scientist. Similarly, the use of the PubMed Related Articles algorithm (PMRA) helps us distinguish between citations coming from within the same subfield, as opposed to outside the subfield. Importantly, this parsing can be implemented for the articles authored by both the treated and the placebo scientists, in a rigorously identical fashion.

Using this information, we create five distinct article subsamples: (1) the set of articles in the top 10% of impact at time of death for each scientist; (2) the set of articles in the bottom 10% of impact at time of death for each scientist; (3) the set of articles in the second and third quartile of the impact distribution at time of death for each scientist; (4) the set of articles in the top 1% of impact at time of death in the PubMed universe; and (5) the set of articles published in a narrow window of three years before the time of death. Note that subsamples one through three use a *relative* benchmark to delineate a set of articles (every scientist in the data must have a top 10% and a bottom 10%, for instance). The fourth subsample uses a *universal* benchmark, and it is possible for scientists in the data to contribute no article to this subsample.

4.2 Main Results

Table 3 presents our core results. Overall, we find the papers of deceased scientists increase in citations slightly after the scientist passes away, but the effect is modest (6.5%) and imprecisely estimated (column 1). Yet, this result conceals heterogenous patterns with respect to the degree of impact these individual pieces of research had achieved by the time of the scientist's death. For the articles that are among the most well-cited in a relative sense (Own Top 10%), the post-death increase in citations is 10% relative to papers of living scientists (column 4), while for the least well-cited articles at the time of death (Own Bottom 10%), the boost is a remarkable 49% (column 2). The papers that lie between the 25% and 75% percentile of citation impact at the time of death (column 3) do not exhibit an effect statistically distinguishable from zero; neither do articles published three years before the author's death nor those that are above the universal benchmark of the 1% most well-cited articles within the *Web of Science.*

Insert Table 3 Here

We also explore the dynamics of the effects uncovered in Table 3. We do so by estimating a specification in which the treatment effect is interacted with a set of indicator variables corresponding to a particular year relative to the scientist's death, and then graphing the

effects and the 95% confidence interval around them (for example, panels A and B of Figure 5 correspond to columns 1 and 2 in Table 4).[13]

Insert Figure 5 Here

Two features of the figure are noteworthy. First, the dynamics amplify the previous results in the sense that we see the effects increasing (in absolute value) monotonically over time: there is no indication that the effects we estimated in Table 3 are merely transitory. Second, there is no discernible evidence of an effect in the years leading up to the death, a finding that validates *ex post* our identification strategy.

We investigate these effects further by breaking down the publications by the age at which the scientist died and the type of death (Table 4 and Figure 5). This reveals that papers written by younger authors (below age 65) receive a 8% boost in citations, while those of older authors do not experience an effect statistically significant from zero (columns 1 and 2, respectively and Figure 5). Those that die young and suddenly receive the largest boost of 14% (column 3). This evidence is consistent with the findings of Lang and Lang which suggested that etchers that die young and suddenly are (much) better remembered than those who are older (Lang and Lang 1988: 93).

Insert Table 4 Here

4.3 Sources of citation increase

One natural question is whether there is something distinctive about the citations that generate the boost documented above. In particular, are these citations more likely to originate from proximate rather than more distant sources, relative to the deceased scientist? We distinguish between three alternative measure of proximity: social, intellectual, and spatial.

To distinguish between socially proximate vs. distant citations, we simply split the citations that accrue to each article in each year between those from articles with an author who

[13] In these specifications, the $AFTER\_DEATH$ term which is common to treated and control publications is also interacted with a complete series of lags and leads relative to the year of death or counterfactual death.

is a former collaborator of the star, versus those where none of the authors had collaborated with the star previously. The intellectual dimension seeks to differentiate between citing articles that belong to the same narrow subfield as the source, versus citing articles that do not belong to the same narrow subfield. Our implementation leverages similarity in keywords, specifically the *PubMed Related Citations Algorithm* (PMRA), described in detail by Azoulay et al. (2016). Finally, we distinguish between geographically proximate vs. distant citers using authors' institutional affiliation obtained from the AAMC Faculty Roster and NIH's CGAF database.

The results for the corresponding specifications are presented in Table 5. Note that the different columns do not correspond to splits of the sample; rather, it is only the dependent variable that changes across specifications. For instance, the first column models the effect of the scientist's passing on the number of citations solely coming from articles who do not include a former collaborator of the deceased (or of the still-living control scientist). Overall, there is little evidence that post-death citations originate relatively more from proximate authors. While the magnitudes are slightly higher for proximate citations, the difference between the effect on proximate vs. non-proximate citations is not itself statistically significant. We tentatively conclude that the citation boost documented in Tables 3 and 4 (as well as Figure 5) reflects a diffuse and increased interest in the deceased's contributions, particularly those that were slightly less well-known while she was alive. (MW Note: This has been changed to make our "diffuse" point more neutral - do you like this?)

Insert Table 5 Here

4.4 Memorialization

These results strongly suggest that the sales force has been mobilized. In clear violation of the efficient citation hypothesis, the perception of scientists' work does in fact change after their death. Our results point to young authors who die suddenly getting a boost in citations, with those who are older or whose death is anticipated experiencing little to no change.

What these results do not explain, however, is why the sales force is being mobilized. As noted in Figure 1, while the boost in citations after death is consistent with the notion of a sales force, by only seeing the boost we can only say that mobilization has occurred. We

cannot say why it has occurred, or even when it is likely to occur. For this reason, further investigation must be done to understand the circumstances that would cause scientists' followers to memorialize them after their death.

To do so, we take a three-step approach. First, we look to measure memorialization to estimate which scientists are most likely to be memorialized. Second, we use that estimation to develop a prediction of posthumous citations based on the extent to which the scientist was memorialized. Finally, we compare our predicted citation rates to the actual citation rates.

4.4.1 Estimating the Determinants of Memorialization

In estimating the amount of memorialization, our goal is to model the number of memory events for scientists at the individual level. To do so, we regress the number of academic memories on the age of the star at death, an indicator variable for whether the death was sudden, and an interaction between these two variables using quasi-maximum likelihood Poisson estimates. In addition, we add controls for the gender of the author, the number of scientists they trained in their career, and the number of distinct collaborators they worked with but whom they did not train. We also include an indicator variable for membership in the National Academy of Sciences, and lastly, their cumulative number of citations at death — two rough proxies for the eminence of these elite scientists within the scientific community.

Table 6 shows our main results. We begin with a simple model including only the controls for gender and sudden death. We then gradually add measures of eminence before concluding with model 8, which includes a control for self-promotion. Columns 2 and 3 show that the number of cumulative citations and publications at death strongly correlate with greater memorialization, while column 4 suggests that no such relationship between funding and memorialization exists. All three measures of eminence are included in column 5, which shows that when combined publications are the strongest predictor of memorialization out of the three. We then add in the logged number of trainees and coauthors in columns 6 and 7, respectively. Both appear to have only a weak relationship to memorialization that is if anything negative. The strongest predictor of memorialization, present in columns 2 through 8, is membership to the National Academy of Sciences.

Insert Table 6 Here

Our main results show evidence of a strong interaction effect for scientists that died young and suddenly, we look to see if a similar interaction effect occurs for memorialization, as Lang and Lang claim (1988). Though the interaction terms for age and sudden deaths are not displayed in Table 6, they can be visualized in Figure 7. Here, we use a model similar to the one described above in column 7 of Table 6, but rather than treating age as a continuous variable we create seven bins for age at the time of death. We then interact the age bins with the indicator for sudden death and graph the coefficients. This reveals that dying suddenly substantially increases the amount of memorialization. While *ceteris paribus* those who die at an older age are memorialized to a greater extent (as is evident in Table 6), for those that die young and suddenly, an interaction occurs which leads them to be significantly more memorialized than their older counterparts. Young scientists which die suddenly accrue nearly three times more memory events than older scientists of a comparable status. Thus, consistent with Lang and Lang (1988), we see evidence that the sudden death of a young scientist mobilizes the sales force to memorialize the deceased scientist.

Insert Figure 7 Here

4.4.2 Who are the Memorializers?

This raises the question of how self-promotion affects the memorialization efforts of the sales force. Though self-promotion is difficult to measure, we look to shed light on this question by including a measure self-promotion in our final specification (column 8) in Table 6. We measure self-promotion using an indicator variable that corresponds to the top quartile of the distribution of unrelated citations as a percentage of self-citations (averaged over each deceased scientist's entire body of work) in our sample. To the extent that this measure is a reasonable proxy, we find that self-promotion is very slightly negatively associated with memorialization, though the relationship is quite noisy and not statistically distinguishable from zero.

To understand the memorialization process further, we look to examine who wrote the "memory events" that comprised the observations in our analysis on memorialization in Sections 4.4.1 above. In examining the 676 deceased scientists in our sample, we identified

1,194 memorializer/deceased pairs. From these pairs, we determine if the author was a co-author or a trainee, as discussed in Section 5.1 For those that were neither, we then coded each of the obituaries written by the remaining memorializers to determine the relationship, where the full text was available on PubMed (nearly 30% of the articles were not available in full). Table 7 shows the relationship breakdown for the memorializer/deceased pairs.

Insert Table 7 Here

Consistent with Lang and Lang (1988), we find that the memorializers did have a social relationship with the decease author, most often as a trainee or collaborator. Over 70% were written by someone the deceased scientist had worked with before (either as a collaborator or as a trainee), while fewer than 10% were from individuals that did not have any relationship at all with the deceased.

4.5 Long-run Citation Afterlife and its Relationship to Memorialization Efforts

The last step of our analysis is to connect the memorialization activity with long-run citation outcomes. We ask whether the memorializers's efforts in the short-run (recall that memory events typically occur within a three-year window after the death) are associated with the long-run citation "afterlife" in our sample of deceased elite scientists. To do so, we regress the predicted and excess citations computed in Section 4.4 on the intensity of memorialization activity, along with a handful of control covariates: gender, degree type, an indicator variable for sudden death, as well as a full set of indicator variables for the scientist's age at the time of death and for his/her calendar year of death. Table 8 reports OLS estimates using the cross-section of deceased scientists for which we were able to ascertain the circumstances of the death (sudden or anticipated). Because memorialization efforts might have a non-linear relationship with long-run citations, we break out the overall count of academic memory events: zero event (185 [27.36%] scientists, the omitted category); exactly one memory event (194 [28.70%] scientists); exactly two memory event (113 [16.72%] scientists); exactly three memory event (54 [11.09%] scientists); exactly four memory event (43 [6.36%] scientists); and five or more memory event (65 [9.76%] scientists).

In Table 8, columns 1a, 1b, and 1c use the log of predicted citations as the outcome variables, whereas columns 2a, 2b, and 2c use the level of excess citations (which can be negative, since it is computed as the difference between actual and predicted citations. Because the distribution of excess citations is both skewed and takes on negative values (see Figure 6), we transform it with a NegLog transformation (Yeo and Johnson 2000), and present results using the transformed outcome in columns 3a, 3b, and 3c.[14] Relative to columns 1a, 2a, and 3a, columns 1b, 2b and 3b slightly modify the data used to compute the outcome variables. In columns 1a, 2a, and 3a, we use all possible citations to build a prediction model for the long-run number of citations for each article published by the deceased scientist. In columns 1b, 2b and 3b, we use the same predictive model but exclude citations from the deceased scientist's memorializers and coauthors. Doing so provides some insight as to the source of increased citations; if the boost in citations is primarily from the scientist's closest friends and colleagues, one would expect these estimates to be small relative to the estimates in columns 1a, 2a, and 3a. Similarly, in columns 1c, 2c, and 3c, we include all citations less those in the three years that immediately follow the death event (but otherwise use the exact same prediction model). The reason to exclude citations that accrue to the deceased scientists' articles in the immediate aftermath of his/her death is that these citations could reflect, at least in part, memorialization efforts (it is not uncommon for obituaries and reminiscences published in scientific journals to have a list of references, for example). By excluding from the count of cumulative citations those that accrue in the period of bereavement, we can be more confident that our measures of predicted and excess citations do not reflect the mechanical impact of memorialization efforts.

Insert Table 8 Here

In Column 1a, we find evidence of a strong positive association between memory events and predicted long-run, posthumous citations. The effect appears monotonic: relative to scientists with no academic memory events, scientists with one memory event are predicted to garner 39% more citations, whereas those in the tail of the memorialization distribution (5 events or more) are predicted to garner 245% more citations. Column 1b shows an almost identical point estimate to column 1a, both measured relatively precisely, suggesting that the boost in citations does not stem from the work of the memorializers and coauthors. Finally, column 1c makes clear this statistical association does not merely reflect awareness by their

[14] $NegLog(x) = log(x)$ if $x > 0$ and $-log(-x)$ if $x < 0$.

immediate community of the turbulent years that immediately follow the passing of these scientists — the estimates are virtually identical to those in columns 1a and 1b.

The estimates presented in columns 2a, 2b, and 2b correspond to the excess citations outcome, entered in levels. These estimates are noisier, with only the coefficient for five or more academic memories entering the model with a statistically significant effect (and a large magnitude, an unexpected boost of 183, 220, and 183 citations at the mean of the data). The other coefficients are all positive, but imprecisely estimated. By using the negative log transform, we produce a less skewed dependent variable. The corresponding estimates, presented in columns 3a, 3b, and 3c, are large in magnitude, precisely estimated, and relatively similar to those displayed in columns 1a, 1b, and 1c.

Overall, the results presented in Table 8 establish the plausibility of memorialization as the underlying mechanism triggering the vibrant "citation afterlife" of deceased scientists. When considered in the context of the results presented in Tables 3-8, our evidence points to the following chain of events: the death of eminent scientists activates a narrow vanguard of colleagues who were proximate to the deceased.[15] It is this vanguard who engages in memorialization efforts, and these efforts in turn bring to the attention of the scientific community at large the work of the deceased, in particular work that may have been overlooked while s/he was alive.

5 Conclusion and Discussion

To recall, the foregoing analysis was motivated by the recognition that science is an especially good setting for advancing our understanding of how reputational entrepreneurship might shape producer legacy. Prior research in the domains of art and politics have presented compelling historical examples where the death of a producer either limits reputational entrepreneurship of the producer/salesman or mobilizes a sales force of people that the producer influenced during their lifetime. The main advantage of the current setting is that there are strong normative and institutional reasons to expect reputational entrepreneurship to be unimportant. In particular, whereas the norms of science imply that a scientist's death should have no impact on the scientific community's valuation of her work, and scientific

[15]Proximity is multidimensional, corresponding to relationships that unfolded in geographic space (such as the case of department or university colleagues), in social space (such as between mentor and trainee, or between coauthors), and in intellectual space (such as shared topics, research questions, and methodologies).

institutions should disseminate and evaluate ideas regardless of efforts to promote such ideas (the efficient citation hypothesis, or ECH), our analysis demonstrates that the promotion of a scientist's work — reputational entrepreneurship — has a significant and lasting impact on its valuation. More specifically, we have seen that the death of a scientist, and especially the sudden death of a young scientist, acts as a catalyst to mobilize a community of scientists seeking to commemorate the deceased. Such commemoration draws attention to the work of the deceased (especially their least-known work), thereby causing the community's valuation of it to be higher.

If the avowed norms of science were fully operative or the mechanisms underlying the scientific marketplace were highly efficient, such commemoration of the deceased would be of no consequence. But this commemoration does matter, thus indicating the weakness of such norms and the inefficiency of the scientific marketplace. In particular, the random event of an untimely death elicits commemoration activity, which increases attention to (and thus higher valuation of) elite scientists' lesser known work. The upshot is that a producer's identity matters in science for the valuation of a given product much as it does in art and politics. This limit on the norm of universalism may reflect the fact that although attention and valuation of scientific work can ignore the identity of the scientist, in practice it is essentially impossible to do so given that key resources — especially jobs and research funding— must be allocated to individual scientists on the basis of their past and prospective bodies of work. Moreover, the social relationships that undergird science are necessarily between individuals (Merton 1970), and we have documented in Table 7 that these social linkages provide the fabric necessary for a narrow vanguard — the "sales force" — to mobilize in the wake of a revered figure's passing.

5.1 Limitations

Before discussing the implications of these findings, it is important to address one clear limitation of our study: that it is restricted to the work of elite academic life scientists. As discussed above, we limit our sample to top academic life scientists in large part because the wealth of information on them allows us to create precise and meaningful counterfactuals. This raises questions as to how this focus affects our analysis and whether our findings would apply to the vast majority of scientists who are lower status.

Some light may be shed on this question by examining the heterogeneity within our sample. As noted above, there is significant variation in status even among the elite scientists. To see if higher status scientists receive a larger boost in citations after their death, we use our main specification described above and split the sample at the median by cumulative publications, citations, and funding at the time of death. The results, displayed in Table 9, do not suggest a much larger effect for high-status scientists. Columns 1 and 2 show that scientists with above the median number of publications receive a boost of 13% while those below the median do not receive a boost statistically different than zero. But, in terms of citations, columns 3 and 4 suggest that if anything, less-cited scientists receive a slightly larger boost, though the estimate is quite noisy. Lastly, those with above median levels of funding are not statistically different than those below the median. Taken together, this could imply that status does not affect the efficacy of reputational entrepreneurship.

Insert Table 9 Here

Yet there are reasons to doubt that we can generalize from an elite sample to the general population of scientists. On the one hand, it is possible that as differences in status become more pronounced, status itself becomes more important. More specifically, the literature on the Matthew Effect (i.e., cumulative advantage processes whereby the attainment of status provides further advantage in attaining even greater status) suggests that a key advantage of high-status scientists is having their work much more widely read (Merton 1968; Azoulay et al. 2014; Simcoe and Waguespack 2011; Cole 1970; Allison et al. 1982). Insofar as this is the case, it implies that the scientists in our sample are more likely to have had their papers read and evaluated than their lower-status peers. The work of elite scientists should thus be relatively insensitive to the benefits of posthumous memorialization. Put differently, while we find that even the highest-status scientists have some work that has been overlooked by the community and is thus sensitive to reputational entrepreneurship, this should *a fortiori* be true for low-status scientists. On the other hand, while the work of lower-status scientists might stand to benefit the most from reputational entrepreneurship, the Matthew Effect would also seem to imply that they are less likely to be the subject of such efforts. Past research suggests that higher-status scientists attract larger numbers of coauthors, research assistants, doctoral students, and admirers (see Zuckerman 1967; Dey et al; 1997; Goldstone 1979; Stewart 1983; Rossiter 1993; Allison and Stewart 1974) — i.e., the "sales force" for the scientist's work. It is also possible that reputational entrepreneurship would be less

valuable for lower-status scientists because audiences will find efforts to promote their work less credible.

In sum, additional research is necessary to clarify how we might generalize beyond our sample of elite scientists. It seems likely that our findings provide a lower bound on the misvaluation that occurs as a result of scientific status and thus the potential returns to reputational entrepreneurship. But they also might provide an upper bound on the level of reputational entrepreneurship that actually occurs given the challenges of mobilizing a (credible) sales force for lower status scientists.

5.2 Implications

The foregoing limitation notwithstanding, our findings have important implications for understanding how scientific works are valued, for understanding how science as a vocation shapes recognition and the allocation of credit, and for reputational entrepreneurship more generally. We now discuss each of these implications in turn.

5.2.1 Citations as Biased Measures of Scientific Achievement

How accurate is the scientific community's valuation of scientific contributions? This question is important because such valuations help to determine which contributions are developed into practical applications, which ideas shape the public's understanding of the natural world, as well as the future trajectory of science (Polanyi 2009 [1966]: 67). Higher valuations of certain ideas and methods encourage scientists and the institutions that support them to redirect their research efforts in some ways but not others (Foster et al. 2015, Mulkay 1972). Citations to prior work are key to this process as they indicate which works, authors, and topics have gained the attention of the community (Merton 1988: 621). Accordingly, citations are often used as a measure of quality for rankings of professors and universities (Altbach 2012; Collyer 2013), and also for the innovativeness of an idea (e.g., Trajtenberg 1990). But the practice of using citations to measure scientific achievement relies on the assumption that they are reliable measures of scientific quality. As long as the Mertonian (1979) norms — universalism, communalism, organized skepticism, and disinterestedness — govern science, this assumption might seem reasonable since the upshot is a community whose members pursue nothing but the truth. However, Merton himself worried that the contributions of high-status scientists gain more credit than is their due (the Matthew Ef-

fect) and the post-Mertonian sociology of science has generally documented the contingent circumstances by which scientific communities shape the production and evaluation of scientific work (Shapin 1982). Moreover, and as reviewed above, the relative absence of arbitrage mechanisms by which a contrarian might take advantage of gaps between scientific valuation and quality makes it unlikely that citations are unbiased indicators of quality.

But while there is reason to doubt that citations are reliable measures of scientific contributions, the nature and extent of the misvaluation is unclear. Building on recent evidence for the Matthew Effect (Azoulay et al., 2014), our analysis demonstrates clear gaps between the implied quality of scientific papers measured by citations and their actual quality. Insofar as this effect centers on the dead scientists' least cited work, the implication is that certain papers do not get the attention they deserve. This is consistent with the presence of significant search costs, as the scientific community struggles to read and evaluate new publications in a timely manner (Gans and Murray 2011; Iaria and Waldinger 2015). Note well though that the effect appears to be permanent. This implies that there is much good science that goes unread and unused. More generally, insofar as our analysis provides clear evidence for how a contingent event can permanently alter the valuation of science as measured by citations, citations must always be taken with a grain of salt.

5.2.2 How Science as a Vocation Shapes Recognition and the Allocation of Credit

Misvaluations arise in part because science struggles to divorce research from the identity of its author. The norms of disinterestedness and universalism belie the fact that science is both a vocation and a means of employment (Merton 1969; Polanyi 2009 [1966]; Gieryn 1983). While the community seeks to evaluate work on its own merit, it must also employ people to teach and to manage laboratories. Similarly, though it publishes papers, science awards prizes and grants to scientists. This creates a circular problem as the community seeks to evaluate works independently from their producers while evaluating producers on the basis of their work. That those who produce the most appreciated research are best suited to review new research exacerbates this problem, creating a strong incentive to factor in the author's identity when evaluating scholarship. This tension between universalism and science as an employment system is most observable in the debate over the "blinding" of the review process; though double-blinded reviews are most common in science, there is significant controversy over the practice precisely because some explicitly wish to use the

author's identity as a signal of quality (Blank 1991; Ceci and Peters 1984). While the salience of identity to the valuation of scientific work is not new in the context of this debate, we demonstrate that even outside of it (or more specifically, after it), the identity of the author materially affects the valuation of scientific work.

This struggle shows science to be nearer to art in its evaluation of work than it would at first appear. There is little debate that the value of a work of art is greatly affected by the identity of the artist. The salience of the artist's identity arises from the fact that art is assessed through the lens of the artist's style (Sgourev and Althuizen 2014; Wohl 2017). For this reason, art is typically organized by artist and reviews are most often done by well-known critics where the identity of both parties is plainly visible. As discussed above, science is structured in stark contrast, organizing its work by subject matter in publications and conferences and emphasizing the objectivity of its review process, all in accordance to the norm of universalism (Merton 1969). Our findings demonstrate that these institutional arrangements are insufficient to completely overcome the incentives created by the employment system within science, however. Just as in the case of Lang and Lang's etchers (1988), the valuation of scientific works is affected by the identity of the author via reputational entrepreneurship.

5.2.3 The Logic of Reputational Entrepreneurship

Finally, our analysis has implications for our understanding of reputational entrepreneurship as a mechanism. Previous work claimed that reputational entrepreneurship significantly impacted valuations but lacked quantifiable measures of valuation and clear counterfactuals (Fine 1996; Lang and Lang 1988). As a result, it was limited in its ability to precisely estimate the extent to which reputational entrepreneurship shifted valuations. Using science as a setting enables us to overcome these obstacles, as it offers citations as a quantifiable measure of value and counterfactual papers to those authored by the dead scientists. This approach yields striking results: reputational entrepreneurship can permanently shift the valuation of work by up to 50% in some cases.

This research design also allows us to shed light on which actors are the most effective reputational entrepreneurs. Prior work tended to focus on either the sales force (e.g., Lang and Lang 1988) or the salesman (e.g., Fine 1996), but did not directly compare the two. Our research design allows for this through the juxtaposition of living scientists and the

memorializers of deceased ones. This comparison reveals the memorializers ("sales force") to be more effective in changing valuations than the scientist herself. Although the question of why exactly this is the case is one we must leave for future research, one likely reason is credibility. Individuals promoting their own work may be limited in their ability to change its valuation because the community recognizes the self-interest at stake and therefore discounts their message. The sales force, by contrast, has little self-interest in commemorating the work of a deceased colleague and therefore their actions are seen as more authentic. As such, the community may be more receptive to their message, and therefore, likely to pay more attention. In this we see an interesting parallel to religion: scholars of religion have noted that major religions frequently begin with the founder's death, which sparks efforts by the disciples to ensure that the founder's life is remembered.[16] It is intriguing that a social phenomenon familiar from the histories of such religions as Christianity, Islam, and Buddhism are also at work in science.

Finally, while prior research on reputational entrepreneurship does not distinguish between shifts in attention and valuation, our results — in particular, that it is the least-cited papers that are most sensitive to reputational entrepreneurship — suggest that attentional processes may be especially important. Our results are not definitive in this regard, nor is it clear to what extent they would generalize to domains beyond science, but they call into question a tendency to assume that reputational entrepreneurship operates by changing the valuations of existing audiences. In bringing overlooked work to the fore, the sales force is able to increase its valuation by changing the sample of work with which the community is engaged (Jerker and Le Mens 2016). That this mechanism is so effective in science, and especially in the work of elite scientists, is testimony to the extent to which search costs inhibit the scientific community's ability to digest new work.

[16]We are grateful to Angela Lee for pointing this to us.

References

Allison, Paul D., and John A. Stewart. 1974. "Productivity Differences Among Scientists: Evidence for Accumulative Advantage." *American Sociological Review* **39**(4): 596-606.

Allison, Paul D., J. Scott Long, and Tad K. Krauze. 1982. "Cumulative Advantage and Inequality in Science." *American Sociological Review* **47**(5): 615-625.

Altbach, Philip G. 2012. "The Globalization of College and University Rankings." *Change: The Magazine of Higher Learning* **44**(1): 26-31.

Azoulay, Pierre, Christian Fons-Rosen, and Joshua S. Graff Zivin. 2015. "Does Science Advance One Funeral at a Time?" NBER Working Paper #21788.

Azoulay, Pierre, Joshua Graff Zivin, and Jialan Wang. 2010. "Superstar Extinction." *Quarterly Journal of Economics* **125**(2): 549-589.

Azoulay, Pierre, Toby Stuart, and Yanbo Wang. 2014. "Matthew: Effect or Fable?" *Management Science* **60**(1): 92-109

Barker, David C., and Christopher Jan Carman. 2010. "Yes WE Can or Yes HE Can? Citizen Preferences Regarding Styles of Representation and Presidential Voting Behavior." *Presidential Studies Quarterly* **40**(3): 431-448.

Blank, Rebecca M. 1991. ″The Effects of Double-Blind Versus Single-Blind Reviewing: Experimental Evidence from the American Economic Review.″ *American Economic Review* **81**(5): 1041-1067.

Botelho, Tristan. 2017. "From Audience to Evaluator: The Effect of Social Influence in Opt-in Evaluation Processes." Working Draft.

Boudreau, Kevin J., Eva C. Guinan, Karim R. Lakhani, and Christoph Riedl. 2016. "Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science." *Management Science* **62**(10): 2765-2783.

Brav, Alon, and James B. Heaton. 2002. "Market Indeterminacy." *Journal of Corporate Law* **28**:518-538.

Bromberg, Minna, and Gary Alan Fine. 2002. "Resurrecting the Red: Pete Seeger and the Purification of Difficult Reputations." *Social Forces* **80**(4): 1135-1155.

Cameron Colin, A. and Pravin K.Trivedi. 1998. *Regression Analysis of Count Data.* New York: Cambridge University Press.

Ceci, Stephen J., and Douglas Peters. 1984. "How Blind is Blind Review?" *American Psychologist* **39**(12): 1491 - 1494.

Cole, Jonathan R., and Stephen Cole. 1972. "The Ortega Hypothesis." *Science* **178**(4059): 368-375.

Cole, Stephen. 1970. "Professional Standing and the Reception of Scientific Discoveries." *American Journal of Sociology* **76**(2): 286-306.

Collyer, Fran. 2013. "The Production of Scholarly Knowledge in the Global Market Arena: University Ranking Systems, Prestige and Power." *Critical Studies in Education* **54**(3): 245-259.

de Solla Price, Derek J. 1963. *Little Science, Big Science.* New York: Columbia University Press.

Denrell, Jerker, and Gaël Le Mens. 2016. "Information Sampling, Belief Synchronization, and Collective Illusions." *Management Science* **63**(2): 528-547.

Dey, Eric L., Jeffrey F. Milem, and Joseph B. Berger. 1997. "Changing Patterns of Publication Productivity: Accumulative Advantage or Institutional Isomorphism?" *Sociology of Education* **70**(4): 308-323.

Evans, James A. 2010. "Industry Collaboration, Scientific Sharing, and the Dissemination of Knowledge." *Social Studies of Science* **40**(5): 757-791.

Fama, Eugene F. 1965. "The Behavior of Stock-Market Prices." *The Journal of Business* **38**(1): 34-105.

Fama, Eugene F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance* **25**(2): 383-417.

Fine, Gary Alan. 1996. "Reputational Entrepreneurs and the Memory of Incompetence: Melting Supporters, Partisan warriors, and Images of President Harding." *American Journal of Sociology* **101**(5): 1159-1193.

Foster, Jacob G., Andrey Rzhetsky, and James A. Evans. 2015. "Tradition and Innovation in Scientists' Research Strategies." *American Sociological Review* **80**(5): 875-908.

Gans, Joshua S., and Fiona Murray, 2011. "Funding Scientific Knowledge: Selection, Disclosure and the Public-private Portfolio." In Joshua Lerner and Scott Stern (Eds.) *The Rate and Direction of Inventive Activity Revisited*, 51-103. Chicago: University of Chicago Press.

Gieryn, Thomas F. 1983. "Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists." *American Sociological Review* **48**(6): 781-795.

Goldstone, Jack A. 1979. "A Deductive Explanation of the Matthew Effect in Science." *Social Studies of Science* **9**(3): 385-391.

Gouriéroux, Christian, Alain Montfort, and Alain Trognon. 1984. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica* **53**(3): 701-707.

Graham, Benjamin. 1984. *The Intelligent Investor*. New York: Harper Collins; Reprint New York: Harper Collins, 2006.

Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica* **52**(4): 909-938.

Henderson, Rebecca, Luigi Orsenigo, and Gary P. Pisano. 1999. "The Pharmaceutical Industry and the Revolution in Molecular Biology: Interactions Among Scientific, Institutional, and 29 Organizational Change." In David C. Mowery, and Richard R. Nelson (Eds.) *Sources of Industrial Leadership*, 267-311. New York: Cambridge University Press.

Holmes, Jack E., and Robert E. Elder Jr. 1989. "Our Best and Worst Presidents: Some Possible Reasons for Perceived Performance." *Presidential Studies Quarterly* **19**(3): 529-557.

Hummel, Patrick. 2010. "Flip-Flopping from Primaries to General Elections." *Journal of Public Economics* **94**(11-12): 1020-1027.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2011. "Multivariate Matching Methods That Are Monotonic Imbalance Bounding." *Journal of the American Statistical Association* **106**(493): 345–361.

Iaria, Alessandro, and Fabian Waldinger. 2015. "International Knowledge Flows: Evidence from the Collapse of International Science in the Wake of WWI." Working Draft.

Jansen, Robert S. 2007. "Resurrection and Appropriation: Reputational Trajectories, Memory Work, and the Political Use of Historical Figures." *American Journal of Sociology* **112**(4): 953-1007.

Jaravel, Xavier, Neviana Petkova, and Alex Bell. 2018. "Team-Specific Capital and Innovation." *American Economic Review* **108**(4-5): 1034-1073.

Kahl, Steven, Young-Kyu Kim, and Damon J. Phillips. 2010. "Identity Sequences and the Early Adoption Pattern of a Jazz Canon, 1920–1929." In Greta Hsu, Giacomo Negro, Özgecan Koçak (Eds.) *Categories in Markets: Origins and Evolution*, Volume 31, 81-113.

Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lang, Gladys Engel, and Kurt Lang. 1988. "Recognition and Renown: The Survival of Artistic Reputation." *American Journal of Sociology* **94**(1): 79-109.

Lewis, David. 1974. "Causation." *The Journal of Philosophy* **70**(17): 556-567.

Leydesdorff, Loet. 1988. "Problems with the 'Measurement' of National Scientific Performance." S*cience and Public Policy* **15**(3):149-152.

Lotka, Alfred J. 1926. "The Frequency Distribution of Scientific Productivity." *Journal of the Washington Academy of Sciences* **16**(12): 317-323.

Malkiel, Burton G. 1989. "Efficient Market Hypothesis." In Peter K. Newman, John Eatwell, and Murray Migate (Eds.) *Finance,* 127-134. London: The New Palgrave.

McCormick, Lisa. 2015. "The Agency of Dead Musicians." *Contemporary Social Science* **10**(3): 323-335.

Merton, Robert K. 1968. "The Matthew Effect in Science." Science **159**(3810): 56-63.

Merton, Robert K. 1970. *On Social Structure and Science*. Chicago: University of Chicago Press.

Merton, Robert K. 1979. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.

Merton, Robert K. 1988. "The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property." Isis **79**(4): 606-623.

Mulkay, Michael Joseph. 1972. *The Social Process of Innovation: A Study in the Sociology of Science*. London: Macmillan.

Murray, Fiona, and Scott Stern. 2007. "Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge?: An Empirical Test of the Anti-Commons Hypothesis." *Journal of Economic Behavior & Organization* **63**(4): 648-687.

Popper, Karl. 1959. *Philosophy of Science*. London: George Allen and Unwin.

Polanyi, Michael. 2009. *The Tacit Dimension*. Chicago: University of Chicago Press; Reprinted Chicago: University of Chicago Press, 2006.

Rossiter, Margaret W. 1993. "The Matthew Matilda Effect in Science." *Social Studies of Science* **23**(2): 325-341.

Segalla, Michael. 2008. "Publishing in the Right Place or Publishing the Right Thing: Journal Targeting and Citations′ Strategies for Promotion and Tenure Committees." *European Journal of International Management* **2**(2): 122-127.

Sgourev, Stoyan V., and Niek Althuizen. 2014. "'Notable' or 'Not Able' When Are Acts of Inconsistency Rewarded?" *American Sociological Review* **79**(2): 282-302.

Shapin, Steven. 1982. "History of Science and its Sociological Reconstructions." *History of Science* **20**(3): 157-211.

Simcoe, Timothy S., and Dave M. Waguespack.2011. "Status, Quality, and Attention: What′s in a (Missing) Name?" *Management Science* **57**(2): 274-290.

Stewart, John A. 1983. "Achievement and Ascriptive Processes in the Recognition of Scientific Articles." *Social Forces* **62**(1): 166-189.

Trajtenberg, Manuel. 1990. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *The RAND Journal of Economics* **21**(1): 172-187.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Benjamin Jones. 2013. "Atypical Combinations and Scientific Impact." *Science* **342**(6157): 468-472.

Wohl, Hannah. 2017. "Signature Styles: The Production of Uniqueness." Working Draft.

Wooldridge, Jeffrey M. 1997. "Quasi-Likelihood Methods for Count Data." In M. Hashem Peseran and Peter Schmidy (Eds.) *Handbook of Applied Econometrics Volume 2: Microeconomics*, 321-368. Oxford: Wiley-Blackwell.

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi. 2007. "The Increasing Dominance of Teams in Production of Knowledge." *Science* **316**(5827): 1036-1039.

Yeo, In-Kwon, and Richard A. Johnson. 2000. "A New Family of Power Transformations to Improve Normality or Symmetry." *Biometrika* **87**(4): 954-959.

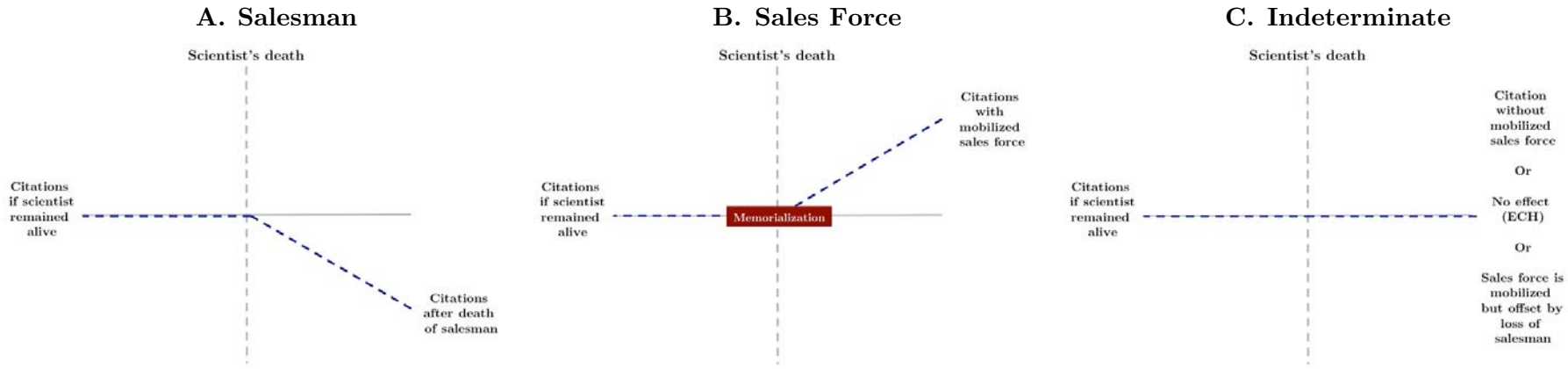
Zuckerman, Ezra W. 2012a. "Construction, Concentration, and (Dis)continuities in Social Valuations." *Annual Review of* Sociology **38**: 223-245.

Zuckerman, Ezra W. 2012b. "Market Efficiency: a Sociological Perspective." In Karen Knorr Cetina and Alex Preda (Eds.) *The Oxford Handbook of the Sociology of Finance*, 223-249. Oxford: Oxford University Press.

Zuckerman, Harriet. 1967. "Nobel Laureates in Science: Patterns of Productivity, Collaboration, and Authorship." *American Sociological Review* **32**(3): 391-403.

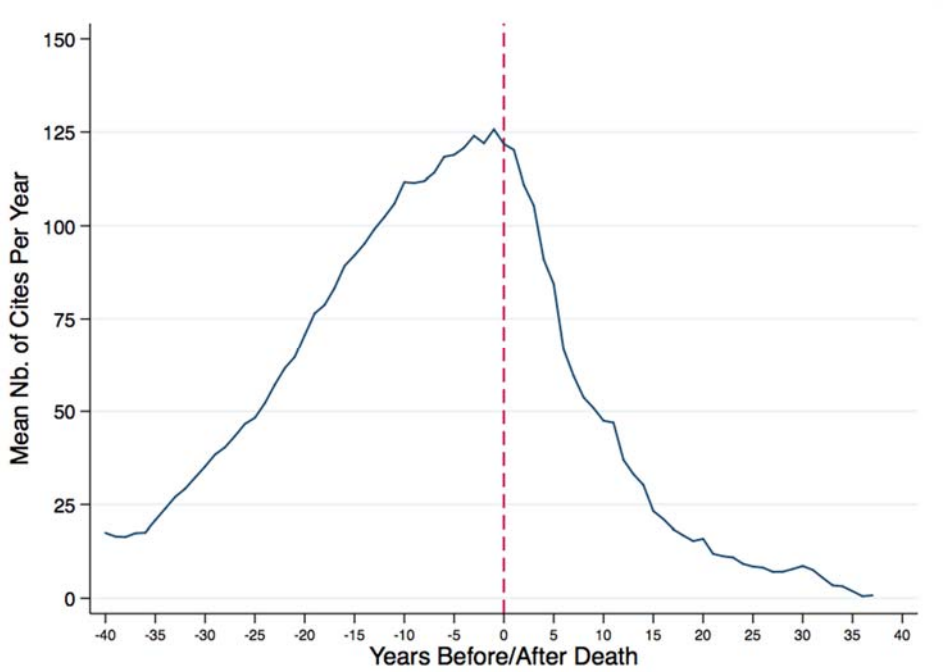
.

Figure 1
Illustration of Different Predicted Results Depending on the Importance of the Salesman and Sales Force



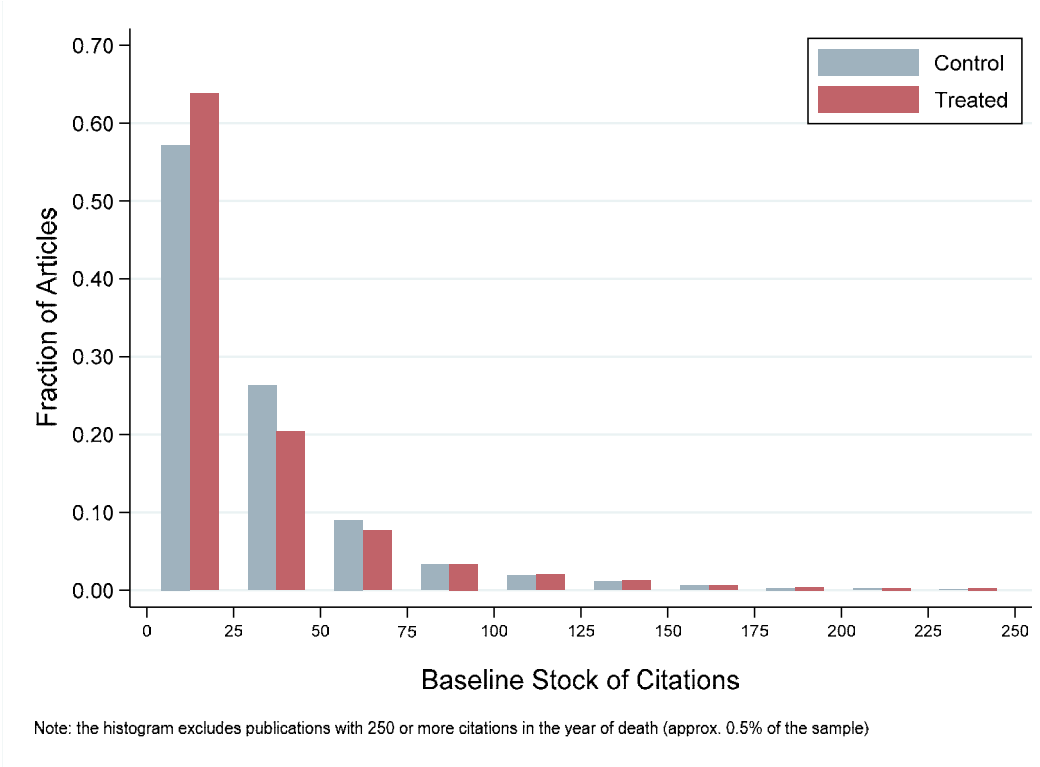
Note: In the above panels, the horizontal gray line through the middle of the graphs represents the cumulative number of citations a given scientist's publication would have accrued had the scientist remained alive. The dashed blue line in Panel A represents the predicted decrease in citations after the scientist's death (represented by the horizontal gray dashed line) under the hypothesis of the salesman as the key reputational entrepreneur. In Panel B, the dashed blue line represents the predicted increase in citations as a result of the mobilization of the sales force. Finally, Panel C displays the third possible outcome in which the effect of death is indeterminant as either there is no effect, consistent with the Efficient Citations hypothesis (ECH), or the effect of the absence of the salesman and the boost from the sales force counteract each other.

Figure 2
Mean Yearly Citations in Relation to Year of Death



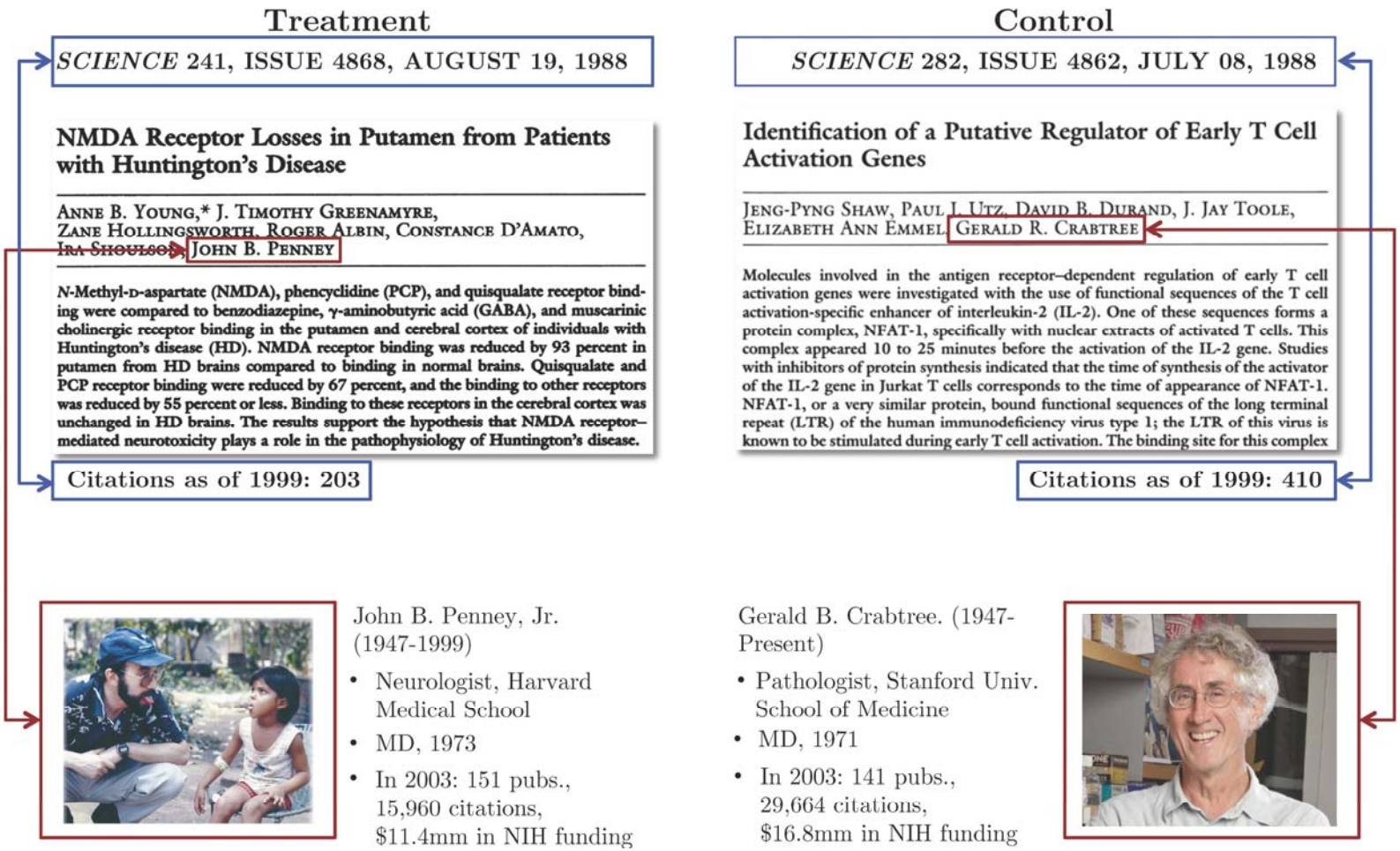
Note: We compute the average total number of citations accrued per year scientist to the 676 treated scientists in the 40 years before and after their death. The dashed-red line indicates the year in which the scientist passes away.

Figure 3
Baseline Stock of Citations at Baseline



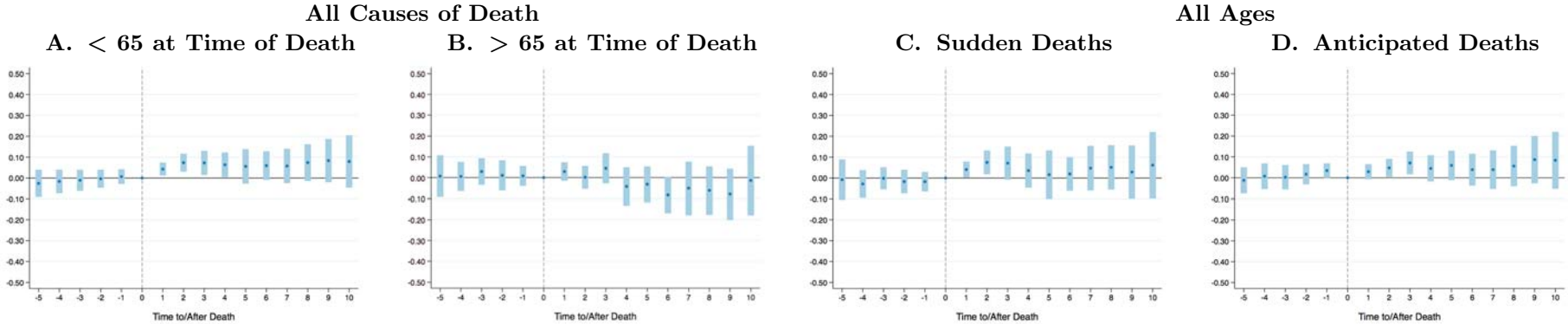
Note: We compute the cumulative number of citations up to the year that immediately precedes the year of death (or the counterfactual year of death) for the 18,896 publications by treated scientists and the 128,218 publications by control scientist.

Figure 4
Matching Procedure to Identify Treatment and Control Articles



Note: The two articles above illustrate the Coarsened Exact Matching (CEM) procedure (Appendix C provides more details). These two articles appeared in the journal Science in 1988. They were both in the same percentile of citations (the top percentile) at the time of death (1999; Young et al., had 203 citations while Shaw et al. had 410 citations). Note that John Penney and Gerald Crabtree are both in last authorship position. They also obtained their MD within a year of each other. This procedure led the Young et al article to be matched with 15 other articles in addition to the Shaw et al article. This same method was followed for the other 18,218 treated articles in our sample.

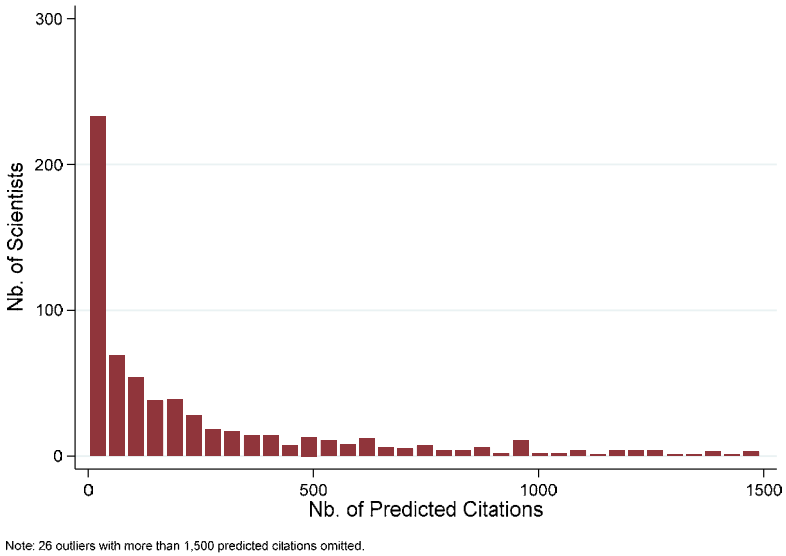
Figure 5
Effect of a Scientist's Death on the Reception of their Work – All Publications



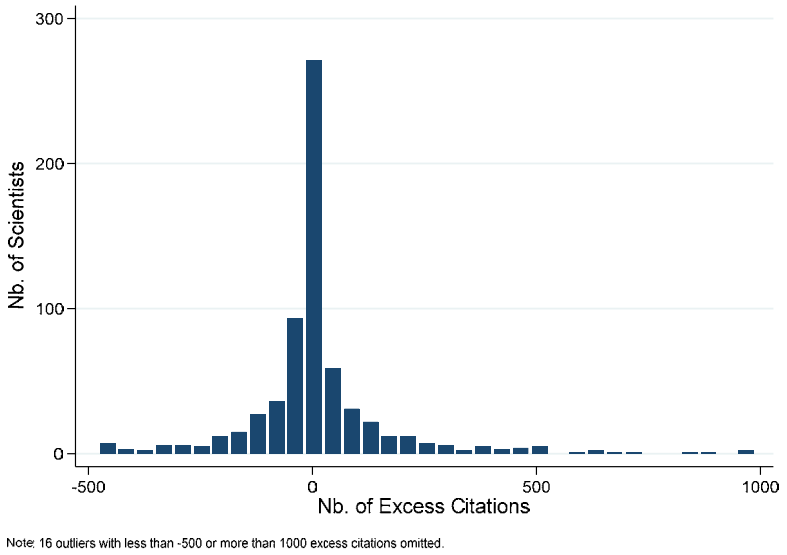
Note: The solid blue dots in the above plots correspond to coefficient estimates stemming from conditional (scientist) fixed effects Poisson specifications in which publication flows are regressed onto year effects, article age effects, as well as 15 interaction terms between treatment status and the number of years before/after the death of the author (the indicator variable for treatment status interacted with the year of death is omitted). The specifications also include a full set of lead and lag terms common to both the treated and control articles to fully account for transitory trends in citations around the time of the death. The 95% confidence interval (corresponding to (QML) robust standard errors, clustered around the scientist) around these estimates is plotted with light blue bars; Panels A and B correspond to dynamic versions of the specification in columns (1-2) of Table 4.

Figure 6
Histograms of Predicted Citations and "Excess" Citations

A. Histogram of Predicted Citations

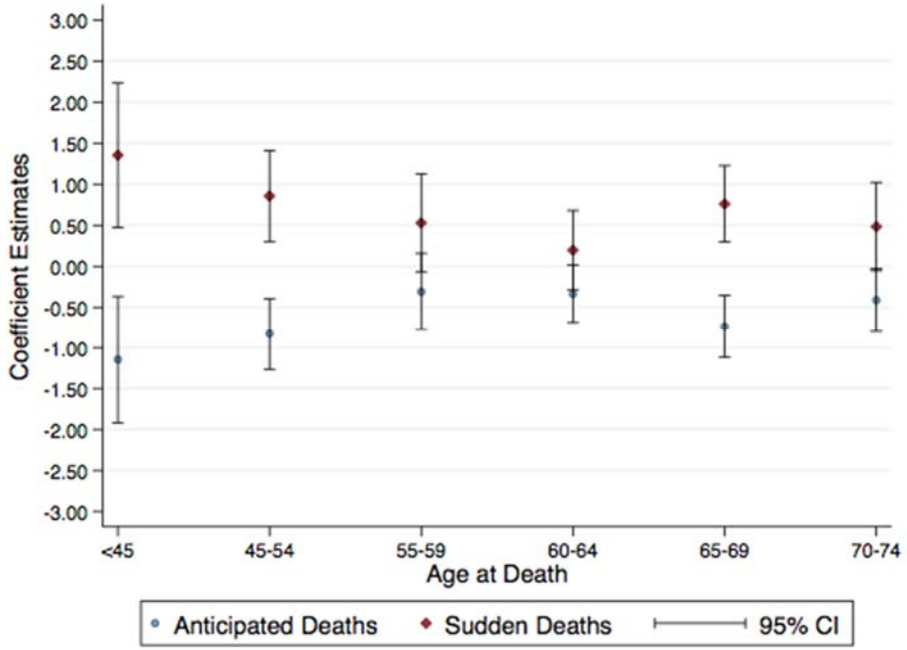


B. Histogram of "Excess" Citations



Note: The blue area in graph A represents the number of citations after death a scientist is predicted to receive in the negative binomial estimates described in section 3.5. This area can only be greater than zero as citations cannot be subtracted. This means that scientists can only gain, not lose, citations posthumously. The red area in graph B represents the difference between the actual cumulative number of citations the scientist received after dying minus the predicted estimate. This amount can be (and is often) negative as scientists receive fewer posthumous citations than our model predicted.

Figure 7
Determinants of Memorialization – Academic Memories



Note: The solid dots in the above plot correspond to coefficient estimates stemming from conditional (scientist) fixed effects Poisson specifications in which the number of memories created for a scientist after their death is regressed onto effects for the scientist's degree type and year of death, as well as six interaction terms between treatment status and the age at which the scientist died. The specifications also includes controls for the scientists gender, number of publications, amount of funding, number or trainees, number of coauthors, and whether or not the scientist was a member of the NAS. The results of a simplified model are presented in Table 7, column 7; the only difference between these this model and the graph above is the use of age as a continuous variable in Table 7 compared to the discrete age bins in the graph above. An academic memory is a festschrift, symposium, obituary in academic journal, or NAS biographical memoirs. The black brackets represent the 95% confidence intervals (calculated using the QML standard errors) and anticipated and sudden deaths are distinguished using blue and red dots, respectively.

Table 1: Summary Statistics of Deceased Scientists (N=676)

	Mean	Median	Std. Dev.	Min.	Max.
Investigator Year of Birth	1927	1927	11.888	1893	1960
Investigator Degree Year	1954	1954	12.505	1921	1988
Investigator Death Year	1990	1992	9.251	1969	2003
Investigator Age at Death	63	64	10.241	35	89
Female	0.087	0	0.282	0	1
MD Degree	0.456	0	0.498	0	1
PhD Degree	0.450	0	0.498	0	1
MD/PhD Degree	0.095	0	0.293	0	1
Investigator Death was Sudden	0.454	0	0.498	0	1
Investigator Death was Anticipated	0.503	1	0.500	0	1
Investigator Cause of Death is Unknown	0.043	0	0.203	0	1
Investigator Cuml. Nb. of Publications	126	102	105	10	1,380
Investigator Cuml. Nb. of Citations	7,044	4,530	7,836	69	72,122
Investigator Cuml. Nb. of Posthumous Predicted Citations	319	109	529	0	4,201
Investigator Cuml. Nb. of Posthumous "Excess" Citations	24	-1	289	-1,170	3,040
Investigator Cuml. Amount of Funding	15,646,305	9,481,224	24,942,410	0	329,968,960
Memorialization Efforts					
Total Nb. Memory Events	4.062	3	4.313	0	61
Total Nb. Academic Memory Events	2.003	1	2.581	0	20
New York Times Obituary	0.324	0	0.471	0	1
Wikipedia Page	0.241	0	0.428	0	1
Named Award	0.226	0	0.419	0	1
Festschrift or Symposium	0.101	0	0.329	0	2

Note: Sample consists of 676 superstar life scientists who died while still actively engaged in research. See Appendix A for more details on the sample construction. Note that 97 (14%) of the treated scientists are NIH intramural scientists and therefore not eligible for NIH funding, resulting in zero cumulative NIH funding.

Table 2: Summary Statistics of Control & Treated Articles at Baseline

	Control Publications (N=128,896)					Treated Publications (N=18,218)				
	Mean	**Med.**	**S.D.**	**Min.**	**Max.**	**Mean**	**Med.**	**S.D.**	**Min.**	**Max.**
Investigator Year of Birth	1930	1930	9.918	1891	1964	1929	1929	10.109	1893	1960
Investigator Degree Year	1957	1957	10.115	1916	1985	1956	1956	10.348	1921	1988
Investigator Death Year	1994	1995	7.371	1969	2003	1993	1995	8.039	1969	2003
Investigator Age at Death	61	61	8.661	35	89	64	64	8.509	35	89
Female	0.050	0	0.218	0	1	0.058	0	0.234	0	1
Investigator Death was Sudden	0.456	0	0.498	0	1	0.425	0	0.494	0	1
Investigator Death was Anticipated	0.514	1	0.500	0	1	0.542	1	0.498	0	1
MD Degree	0.366	0	0.482	0	1	0.415	0	0.493	0	1
PhD Degree	0.551	1	0.497	0	1	0.446	0	0.497	0	1
MD/PhD Degree	0.083	0	0.276	0	1	0.140	0	0.347	0	1
Nb. Authors Per Publication	3.095	3	1.255	1	15	3.188	3	1.481	1	15
Publication Age in Year of Death	3.897	4	2.369	0	9	4.404	4	2.522	0	9
Year Article was Written	1994	1995	7.371	1969	2003	1994	1995	8.039	1969	2003
Total Article Citations	32.494	21	61.908	0	11,505	31.842	16	61.910	0	3,117
Citations by Non-Collaborators	30.074	19	59.243	0	11,295	29.822	15	58.846	0	3,071
Citations by Collaborators	2.420	1	4.852	0	210	2.020	0	4.972	0	184
Citations outside of Field	27.559	17	59.212	0	11,464	27.853	13	58.844	0	3,090
Citations within Field	4.934	3	6.238	0	212	3.989	2	6.209	0	138
Citations from Non-Colocated Field	31.659	20	60.753	0	11505	30.991	16	60.003	0	3070
Citations from Colocated Authors	0.835	0	2.475	0	167	0.851	0	2.932	0	127
Investigator Cuml. Nb. of Publications	195	158	138	7	1,124	207	166	163	10	1,380
Investigator Cuml. Nb. of Citations	14,863	10,282	15,058	17	157,581	13,242	9,780	11,229	69	72,122
Investigator Cuml. Amount of Funding (in millions)	$23.586	$16.355	$25.756	$0	$408.427	$24.137	$14.637	$38.823	$0	$329.969

Note: The sample consists of all of the publications for the 12,935 scientists in the wider sample. See Appendix C for details on the matching procedure. All time-varying covariates are measured in the year of the scientist's death. All cumulative statistics are taken as of the baseline year, the year of death for the treated scientist (and the counterfactual year of death for the control scientist).

Table 3: Main Effect of Scientist's Death – All Ages, All Causes of Death

	Type of Publications					
	All Publications	**Own Bottom 10%**	**Own 25%-75%**	**Own Top 10%**	**Universe Top 1%**	**3 Years Before Death**
After Death	0.063*	0.399***	0.028	0.096*	0.015	0.074
	(0.033)	(0.083)	(0.031)	(0.055)	(0.076)	(0.062)
Nb. of Investigators	6,782	1,556	6,225	2,776	1,066	3,680
Nb. of Source Articles	145,413	2,871	93,314	10,050	2,917	14,925
Nb. of Source Artcl.-Year Obs.	3,927,912	72,409	2,486,563	293,405	75,411	255,831
Log Likelihood	-5,068,850	-39,259	-2,945,129	-713,856	-325,431	-393,223

Note: Estimates stem from fixed effects Poisson specifications. The dependent variable is the total number of citations accrued to a publication in a particular year. All models incorporate a full suite of year effects and article age effects, as well as a term common to both treated and control articles that switches from zero to one after the death of the scientist, to address the concern that age, year and individual fixed effects may not fully account for trends in citations after death. Own Top 10% represents the scientists' top publications measured in citations at the time of death; Own Bottom 10% represents the scientists' lowest publications measured in citations at the time of death, while Own 25%-75% represents the scientists' middle papers by citations. Exponentiating the coefficients and differencing from one yield numbers interpretable as elasticities. For example, the estimates in column (1) imply that the papers deceased scientists see an increase in the number of citations posthumously relative to papers whose author remained alive by $100\times(\exp[0.063]\text{-}1)=6.50\%$. The number of observations varies slightly across columns because the conditional fixed effects specification drops observations corresponding to articles for which there is no variation in activity over the entire observation period. This is also true for the results reported in Tables 3 through 6.

Robust (QML) standard errors in parentheses, clustered at the level of the star scientist. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Effect of Scientist's Death – All Publications

	All Causes of Death		Sudden Deaths		Anticipated Deaths	
	< 65 at Time of Death	**> 65 at Time of Death**	**< 65 at Time of Death**	**> 65 at Time of Death**	**< 65 at Time of Death**	**> 65 at Time of Death**
After Death	0.078**	0.025	0.128**	-0.055	0.046	0.118
	(0.037)	(0.063)	(0.055)	(0.065)	(0.049)	(0.096)
Nb. of Investigators	6,365	2,672	5,306	1,939	4,722	2,023
Nb. of Source Articles	100,496	44,917	47,798	17,891	50,333	24,940
Nb. of Source Artcl.-Year Obs.	2,545,626	1,382,286	1,123,255	543,563	1,333,214	760,684
Log Likelihood	-3,439,023	-1,628,138	-1,534,447	-617,523	-1,824,110	-934,761

Note: Estimates stem from fixed effects Poisson specifications. The dependent variable is the total number of citations accrued to a publication in a particular year. All models incorporate a full suite of year effects and article age effects, as well as a term common to both treated and control articles that switches from zero to one after the death of the scientist, to address the concern that age, year and individual fixed effects may not fully account for trends in citations after death. Exponentiating the coefficients and differencing from one yield numbers interpretable as elasticities. For example, the estimates in column (1) imply that the papers of young scientists see an increase in the number of citations posthumously relative to papers whose author remained alive by a statistically significant $100\times(\exp[0.078]-1)=8.11\%$. The number of observations varies slightly across columns because the conditional fixed effects specification drops observations corresponding to articles for which there is no variation in activity over the entire observation period. This is also true for the results reported in Tables 3 through 6.

Robust (QML) standard errors in parentheses, clustered at the level of the star scientist. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5: Effect of Scientist's Death – Citation Sources

	Coauthor		Field		Location	
	Non-Coauth. Cites	**Coauth. Cites**	**Out-of-Field Cites**	**In-Field Cites**	**Non-Coloc. Cites**	**Coloc. Cites**
After Death	0.044	0.060	0.055	0.089**	0.047	0.143**
	(0.038)	(0.068)	(0.034)	(0.043)	(0.042)	(0.068)
Nb. of Investigators	5,778	5,778	6,502	6,502	5,022	5,022
Nb. of Source Articles	90,722	90,722	124,828	124,828	55,464	55,464
Nb. of Source Artcl.-Year Obs.	2,367,869	2,367,869	3,264,980	3,264,980	1,456,100	1,456,100
Log Likelihood	-3,468,945	-855,386	-4,229,584	-1,437,250	-2,413,017	-354,119

Note: Estimates stem from fixed effects Poisson specifications. The dependent variable is the total number of citations accrued to a publication in a particular year. All models incorporate a full suite of year effects and article age effects, as well as a term common to both treated and control articles that switches from zero to one after the death of the scientist, to address the concern that age, year and individual fixed effects may not fully account for trends in citations after death. Exponentiating the coefficients and differencing from one yield numbers interpretable as elasticities. For example, the estimates in column (6) imply that the papers of deceased scientists see a posthumous increase in the number of citations from collocated scientists by a statistically significant $100\times(\exp[0.146]-1)=15.72\%$. The number of observations varies slightly across columns because the conditional fixed effects specification drops observations corresponding to articles for which there is no variation in activity over the entire observation period. This is also true for the results reported in Tables 3 through 6.

Robust (QML) standard errors in parentheses, clustered at the level of the star scientist. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: Estimating the Determinants of Memorialization – Academic Memories

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(cmltv. citations at death)		0.269***			0.012			
		(0.042)			(0.069)			
Ln(cmltv. publications at death)			0.513***		0.497***	0.544***	0.647***	0.608***
			(0.053)		(0.089)	(0.064)	(0.096)	(0.107)
Ln(cmltv. funding at death)				0.019	-0.000			
				(0.012)	(0.010)			
Member of the NAS		0.591***	0.661***	0.827***	0.651***	0.686***	0.635***	0.632***
		(0.092)	(0.084)	(0.088)	(0.090)	(0.087)	(0.087)	(0.100)
Ln(Nb. of past trainees)						-0.050		-0.033
						(0.064)		(0.067)
Ln(Nb. of past coauthors [non-trainees])							-0.161*	-0.109
							(0.093)	(0.092)
Self-Promoter								-0.020
								(0.166)
Female	-0.354**	-0.185	-0.090	-0.295*	-0.093	-0.094	-0.070	-0.074
	(0.166)	(0.149)	(0.145)	(0.157)	(0.144)	(0.145)	(0.144)	(0.145)
Death is Sudden	0.169*	0.191**	0.206**	0.162*	0.208**	0.202**	0.204**	0.204**
	(0.094)	(0.086)	(0.084)	(0.088)	(0.083)	(0.085)	(0.084)	(0.084)
Nb. of Scientists	647	647	647	647	647	647	647	647
Pseudo-R^2	0.087	0.171	0.191	0.146	0.192	0.192	0.193	0.195

Note: Estimates stem from Poisson specifications. The dependent variable is the total number of academic memories created for a scientist posthumously. An academic memory is a festschrift, symposium, obituary in an academic journal, or NAS memoir. All models include controls for degree type, death year, six age bins as well as the interaction terms for each age bin and sudden death. Self-Promoter is an indicator variable that corresponds to the top quartile of the distribution of unrelated citations as a percentage of self-citations (averaged over each deceased scientist's entire body of work) in our sample. Additionally, the specification in column (8) includes as a control the fraction of *related* self-citations as proportion of all self-citation (similarly averaged over the entire body of work of the focal scientist), to account for his/her propensity to cite his/her own work more generally. The sample size of 647 consists of the 676 original scientists less 29 scientists for which we were unable to confirm the cause of death (anticipated or sudden).

Robust [QML] standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7: Memorializers' Relationships to Deceased Elite Scientists

Type of Relationship	Specific Connection	Percentage of Sample
Social	Trainee	45.49%
	Collaborator	24.11%
	Family	0.48%
	Trained together	0.48%
		70.56%
Intellectual	Colleague in same field	12.47%
	Journal editor	0.59%
		16.14%
None	No social relation	6.18%
	Historian	1.54%
	Journalist	1.54%
		9.26%
Geographic	Shared employer	7.13%
		7.13%

Note: The sample is as a percentage of the 842 memorializer-deceased pairs for which information was available from PubMed. While this meant that we could not account for 352, or 20% of the pairs, there is no reason to suggest that they would be systematically different than above. All categories are mutually exclusive by design.

Table 8: Long-run Citation Afterlife and its Relationship to Memorialization Efforts

	Ln(Predicted Citations)			Excess Citations			NegLn(Excess Citations)		
	All citations	Excl. mems. and coauthors	Excl. 3 years post-death	All citations	Excl. mems. and coauthors	Excl. 3 years post-death	All citations	Excl. mems. and coauthors	Excl. 3 years post-death
Scientists w/ 1 Acad. Memory Event	0.336*	0.334*	0.335*	12.573	17.571	12.576	0.788*	0.681*	0.666*
	(0.175)	(0.174)	(0.174)	(28.935)	(29.917)	(28.937)	(0.407)	(0.407)	(0.390)
Scientists w/ 2 Acad. Memory Events	0.594***	0.586***	0.593***	36.295	47.493	36.303	1.342***	1.378***	1.080**
	(0.204)	(0.202)	(0.202)	(32.732)	(33.579)	(32.734)	(0.487)	(0.477)	(0.476)
Scientists w/ 3 Acad. Memory Events	0.636***	0.628***	0.632***	32.906	51.617	32.907	1.929***	1.750***	1.571***
	(0.230)	(0.229)	(0.229)	(32.953)	(33.329)	(32.956)	(0.557)	(0.553)	(0.534)
Scientists w/ 4 Acad. Memory Events	0.887***	0.867***	0.888***	42.397	68.708	42.399	1.041	1.233*	0.773
	(0.267)	(0.266)	(0.265)	(46.220)	(45.063)	(46.219)	(0.702)	(0.678)	(0.681)
Scientists w/ 5+ Acad. Memory Events	1.239***	1.214***	1.234***	182.623**	219.806***	182.666**	2.035***	2.006***	1.762***
	(0.276)	(0.275)	(0.274)	(78.170)	(82.153)	(78.179)	(0.604)	(0.593)	(0.570)
Female	-0.699**	-0.688**	-0.698**	-42.390	-52.807*	-42.394	-0.624	-0.727	-0.520
	(0.283)	(0.281)	(0.281)	(27.665)	(27.435)	(27.671)	(0.560)	(0.543)	(0.539)
Death is Sudden	-0.223*	-0.226*	-0.222*	-39.411	-44.719*	-39.408	-0.585*	-0.662**	-0.643**
	(0.135)	(0.134)	(0.134)	(25.880)	(26.763)	(25.882)	(0.319)	(0.314)	(0.306)
Constant	1.639***	1.478**	1.449**	-87.884	-86.384	-87.888	-0.586	-0.720	-1.376
	(0.617)	(0.623)	(0.616)	(73.016)	(74.392)	(73.018)	(1.294)	(1.196)	(1.224)
Nb. Treated Scientists	647	647	647	647	647	647	647	647	647
Adjusted R^2	0.323	0.323	0.375	0.045	0.058	0.045	0.040	0.046	0.030

Note: Estimates stem from OLS specifications. The dependent variable in columns 1a, 1b, and 1c is the log of the number of *predicted* long-run, posthumous citations (based on the prediction model presented in Section 3.5). Column 1a includes all citations, while column 1b subtracts out citations by memorializers and coauthors. Finally, columns 1c includes all citations less those in the three years immediately after the scientist's death. The dependent variables in columns 2a, 2b, and 2c is the number of "excess" citations (which is simply the number of actual posthumous citations minus the number of predicted posthumous citations), also broken down by all citations, all citations less memorializers and coauthors, and all citations minus the immediate three years after the scientist's death. Because the distribution of excess citations is both skewed and takes on negative values, columns 3a, 3b, and 3c perform a NegLog transformation (Yeo and Johnson 2000) before estimating the model, using the same breakdown all citations, all citations less memorializers and coauthors, and all citations minus the immediate three years after the scientist's death. All models include controls for the age at death, year of death, and degree type. The sample size of 647 consists of the 676 original scientists less 29 scientists for which we were unable to confirm the cause of death (anticipated or sudden).

Robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 9: Effect of Scientist's Death – Scientist's Status

	Publications		Citations		Funding	
	Below Median	**Above Median**	**Below Median**	**Above Median**	**Below Median**	**Above Median**
After Death	0.015	0.122**	0.072*	0.058	0.047	0.077
	(0.039)	(0.053)	(0.038)	(0.049)	(0.043)	(0.056)
Nb. of Investigators	5,800	2,200	5,455	2,238	5,066	2,439
Nb. of Source Articles	72,973	72,440	72,707	72,706	67,001	66,991
Nb. of Source Artcl.-Year Obs.	1,977,527	1,950,385	2,024,864	1,903,048	1,788,996	1,838,386
Log Likelihood	-2,534,940	-2,529,819	-2,316,243	-2,751,326	-2,268,927	-2,398,501

Note: Estimates stem from fixed effects Poisson specifications. The dependent variable is the total number of citations accrued to a publication in a particular year. All models incorporate a full suite of year effects and article age effects, as well as a term common to both treated and control articles that switches from zero to one after the death of the scientist, to address the concern that age, year and individual fixed effects may not fully account for trends in citations after death. Exponentiating the coefficients and differencing from one yield numbers interpretable as elasticities. For example, the estimates in column (1) imply that the papers of scientists with below the median number of cumulative publication at the time of their death see an increase in the number of citations posthumously relative to papers whose author remained alive by a statistically significant $100\times(\exp[0.015]-1)=1.51\%$. The number of observations varies slightly across columns because the conditional fixed effects specification drops observations corresponding to articles for which there is no variation in activity over the entire observation period. This is also true for the results reported in Tables 3 through 6.

Robust (QML) standard errors in parentheses, clustered at the level of the star scientist. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Supplementary Online Material

Appendix A:
Criteria for Delineating the Set of 12,935 "Superstars"

Highly Funded Scientists. Our first data source is the Consolidated Grant/Applicant File (CGAF) from the U.S. National Institutes of Health (NIH). This dataset records information about grants awarded to extramural researchers funded by the NIH since 1938. Using the CGAF and focusing only on direct costs associated with research grants, we compute individual cumulative totals for the decades 1977-1986, 1987-1996, and 1997-2006, deflating the earlier years by the Biomedical Research Producer Price Index. We also recompute these totals excluding large center grants that usually fund groups of investigators (M01 and P01 grants). Scientists whose totals lie above the 95^{th} percentile of either distribution constitute our first group of superstars. In this group, the least well-funded investigator garnered $10.5 million in career NIH funding and the most well-funded $462.6 million.[i]

Highly Cited Scientists. Despite the preeminent role of the NIH in the funding of public biomedical research, the above indicator of "superstardom" biases the sample towards scientists conducting relatively expensive research. We complement this first group with a second composed of highly cited scientists identified by the Institute for Scientific Information. A Highly Cited listing means that an individual was among the 250 most cited researchers for their published articles between 1981 and 1999, within a broad scientific field.[ii]

Top Patenters. We add to these groups academic life scientists who belong in the top percentile of the patent distribution among academics—those who were granted 17 patents or more between 1976 and 2004.

Members of the National Academy of Science and of the Institute of Medicine. We add to these groups academic life scientists who were elected to the National Academy of Science or the Institute of Medicine between 1970 and 2013.

MERIT Awardees of the NIH. Initiated in the mid-1980s, the MERIT Award program extends funding for up to 5 years (but typically 3 years) to a select number of NIH-funded investigators *"who have demonstrated superior competence, outstanding productivity during their previous research endeavors and are leaders in their field with paradigm-shifting ideas."* The specific details governing selection vary across the component institutes of the NIH, but the essential feature of the program is that only researchers holding an R01 grant in its second or later cycle are eligible. Further, the application must be scored in the top percentile in a given funding cycle.

Former and current Howard Hughes Medical Investigators (HHMIs). Every three years, the Howard Hughes Medical Institute selects a small cohort of mid-career biomedical scientists with the potential to revolutionize their respective subfields. Once selected, HHMIs continue to be based at their institutions, typically leading a research group of 10 to 25 students, postdoctoral associates and technicians. Their appointment is reviewed every five years, based solely on their most important contributions during the cycle.[iii]

Early career prize winners. We also included winners of the Pew, Searle, Beckman, Rita Allen, and Packard scholarships for the years 1981 through 2000. Every year, these charitable foundations provide seed

[i] We perform a similar exercise for scientists employed by the intramural campus of the NIH. These scientists are not eligible to receive extramural funds, but the NIH keeps records of the number of "internal projects" each intramural scientist leads. We include in the elite sample the top five percentiles of intramural scientists according to this metric.

[ii] The relevant scientific fields in the life sciences are microbiology, biochemistry, psychiatry/psychology, neuroscience, molecular biology & genetics, immunology, pharmacology, and clinical medicine.

[iii] See Azoulay et al. (2011) for more details and an evaluation of this program.

funding to between 20 and 40 young academic life scientists. These scholarships are the most prestigious accolades that young researchers can receive in the first two years of their careers as independent investigators.

Appendix B: Linking Scientists with their Journal Articles

The source of our publication data is *PubMed*, a bibliographic database maintained by the U.S. National Library of Medicine that is searchable on the web at no cost.[iv] *PubMed* contains over 14 million citations from 4,800 journals published in the United States and more than 70 other countries from 1950 to the present. The subject scope of this database is biomedicine and health, broadly defined to encompass those areas of the life sciences, behavioral sciences, chemical sciences, and bioengineering that inform research in health-related fields. In order to effectively mine this publicly-available data source, we designed PubHarvester, an open-source software tool that automates the process of gathering publication information for individual life scientists (see Azoulay et al. 2006 for a complete description of the software). PubHarvester is fast, simple to use, and reliable. Its output consists of a series of reports that can be easily imported by statistical software packages.

This software tool does not obviate the two challenges faced by empirical researchers when attempting to accurately link individual scientists with their published output. The first relates to what one might term "Type I Error," whereby we mistakenly attribute to a scientist a journal article actually authored by a namesake; The second relates to "Type II error," whereby we conservatively exclude from a scientist's publication roster legitimate articles:

Namesakes and popular names. *PubMed* does not assign unique identifiers to the authors of the publications they index. They identify authors simply by their last name, up to two initials, and an optional suffix. This makes it difficult to unambiguously assign publication output to individual scientists, especially when their last name is relatively common.

Inconsistent publication names. The opposite danger, that of recording too few publications, also looms large, since scientists are often inconsistent in the choice of names they choose to publish under. By far the most common source of error is the haphazard use of a middle initial. Other errors stem from inconsistent use of suffixes (Jr., Sr., 2nd, etc.), or from multiple patronyms due to changes in spousal status.

To deal with these serious measurement problems, we opted for a labor-intensive approach: the design of individual search queries that relies on relevant scientific keywords, the names of frequent collaborators, journal names, as well as institutional affiliations. We are aided in the time-consuming process of query design by the availability of a reliable archival data source, namely, these scientists' CVs and biosketches. PubHarvester provides the option to use such custom queries in lieu of a completely generic query (e.g, `"azoulay p"[au]` or `"graff zivin js"[au]`). As an example, one can examine the publications of Scott A. Waldman, an eminent pharmacologist located in Philadelphia, PA at Thomas Jefferson University. Waldman is a relatively frequent name in the United States (with 208 researchers with an identical patronym in the AAMC faculty roster); the combination `"waldman s"` is common to 3 researchers in the same database. A simple search query for `"waldman sa"[au] OR "waldman s"[au]` returns 377 publications at the time of this writing. However, a more refined query, based on Professor Waldman's biosketch returns only 256 publications.[v]

The above example also makes clear how we deal with the issue of inconsistent publication names. PubHarvester gives the end-user the option to choose up to four *PubMed*-formatted names under which publications can be found for a given researcher. For example, Louis J. Tobian, Jr. publishes under `"tobian l"`, `"tobian l jr"`, and `"tobian lj"`, and all three names need to be provided as inputs to generate a complete publication listing. Furthermore, even though Tobian is a relatively rare name, the search query needs to be modified to account for these name variations, as in `("tobian l"[au] OR "tobian lj"[au])`.

[iv] `http://www.pubmed.gov/`

[v] `((((("waldman sa"[au] NOT (ether OR anesthesia)) OR ("waldman s"[au] AND (murad OR philadelphia[ad] OR west point[ad] OR wong p[au] OR lasseter kc[au] OR colorectal))) AND 1980:2013[dp])`

Appendix C: Construction of the Control Group

We detail the procedure implemented to identify the control publications that help pin down the life-cycle and secular time effects in our difference-in-differences (DD) specification. Happenstance might yield a sample of publications from aging scientists, or in out-of-fashion fields. More plausibly, citation trends might be subject to idiosyncratic life-cycle patterns, with citation the rates of articles reflecting the trends of the age of the article, the age of the scientist, and the age of the field and methods. Relying solely on publications treated earlier or later as an implicit control group raises the worry that these time-varying omitted variables will not be fully captured by publication age controls.
To address this concern, we create an additional level of difference by selecting control publications. Recall that using the PubMed database we can accurately identify the complete publication history of all the scientists in our sample. From this, the key is to identify articles which are similar to those written by the deceased scientists. Practically, we must recruit control source articles from the set of articles authored by scientists who do not die prematurely. But what makes a satisfactory control group? It is important to distinguish between *ex ante* vs. *ex post* criteria. *Ex ante*, one would like control source articles to have the following properties:

1. to be published contemporaneously with the article from the treated scientist;
2. to be unrelated (in both an intellectual and a social sense) to the article from the treated scientist;
3. to be of similar expected impact and fruitfulness, relative to the article from the treated scientist;
4. to have a similar number of authors as the article from the treated scientist;
5. to have an author in the same authorship position and of approximately the same age as that occupied by the deceased scientist on the authorship roster of the article from the treated scientist.

Ex post, it will be important for the control publications to satisfy an additional condition: the treated and control publications should exhibit very similar trends in publication activity up to the year of treatment (i.e., the year of death for the treated scientist).

Coarsened Exact Matching. To meet these goals, we implement a "Coarsened Exact Matching" (CEM) procedure (Blackwell et al. 2009). The first step is to select a relatively small set of covariates on which we need to guarantee balance *ex ante*. This choice entails judgement, but is strongly guided by the set of criteria listed above. The second step is to create a large number of strata to cover the entire support of the joint distribution of the covariates selected in the previous step. In a third step, each observation is allocated to a unique strata, and for each observation in the treated group, control observations are selected from the same strata.

The procedure is coarse because we do not attempt to precisely match on covariate values; rather, we coarsen the support of the joint distribution of the covariates into a finite number of strata, and we match a treated observation if and only if a control observation can be recruited from this strata. An important advantage of CEM is that the analyst can guarantee the degree of covariate balance *ex ante*, but this comes at a cost: the more fine-grained the partition of the support for the joint distribution (i.e., the higher the number of strata), the larger the number of unmatched treated observations.

Implementation. We identify controls based on the following set of covariates (t denotes the year of death): scientist career age; citations received by the article up to year t; number of authors; position of the star author on the authorship roster (only last authorship position is considered); journal; and year of publication. The first three covariates only need to match within relatively coarse bins. For instance, we create nine career age categories: less than 10 years; between 10 and 20 years; between 20 and 25 years; between 25 and 30 years; between 30 and 35 years; between 35 and 40 years; between 40 and 45 years; between 45 and 50 years, over 50 years of career age. Similarly, we coarsen the distribution of citations at baseline into five mutually exclusive bins: zero citations; between one and 10 citations; between 10 and 50

citations; between 50 and 120 citations; and more than 120 citations. In contrast, we impose an exact match on journal, publication year, and the star's authorship position.

We match approximately 45.57% of the treated source articles in this way. Some further trimming of the control articles is needed. First, we eliminate any article with more than 15 coauthors. Second, we remove any article in a non-English speaking journal. Third, we drop any control that shares any author with the treated source. Finally, we drop from the data any source article that finds itself an orphan (i.e., not paired with any control) at the conclusion of this process. Figure 4 provides an illustrative example.

The final sample has 18,523 treated source articles and 128,591 control source articles. As can be seen in Figure 3, the distribution of citations, measured up to the baseline year, is very similar between treated and control publications. As well, there is no evidence of preexisting trends in activity, as demonstrated by the coefficient estimates graphed in Figure 5. In Table 2, treated and control publications are very well-balanced on the covariates that formed the basis of the CEM matching procedure. This is true almost by construction. What is more surprising (and also welcome) is that the procedure balances a number of covariates that were not used as inputs for matching, such as various metrics of scientist eminence. For other covariates, we can detect statistically significant mean differences, though they do not appear to be substantively meaningful (e.g., 14.0% of control stars vs. 8.5% of treated stars have both an MD and PhD).

Sensitivity Analyses. Human judgment matters for the outcome of the CEM procedure insofar as one must draw a list of "reasonable" covariates to match on, as well as decide on the degree of coarsening to impose. We have verified that slight variations in the implementation (e.g., varying slightly the number of cutoff points for the stock of baseline citations for the source; focusing on birth age as opposed to career age for the stars) have little impact on the main results.

Appendix D: Memorialization Data Collection

We started the data collection for memorialization by searching for any academic publications about the deceased scientists in PubMed. We identified these articles by searching for the authors last name in the title of the article (e.g, `wahlen[ti]`). We then reviewed and classified these articles as one of several mutually exclusive type: academic article, festschrift, symposium, or National Academy of Science Memoir. These together comprise the memory events we label academic memories. As noted in the table below, we found on average two academic memories per scientist. There is a significant skew, however, with the median scientist only receiving one while the maximum recorded was 20 (for physician Solomon Berson).

To get a broader view of memorialization, we then repeated the same process using Google searches. We searched for articles, web sites, and obituaries by searching for the scientists name, field, and death year (e.g, `wahlen AND physician AND 1987`). We categorized these findings as university web posts, New York Times obituaries, other newspaper obituaries, Wikipedia pages, and miscellaneous online obituaries. We labeled these memories "popular memories," and we again found on average just over two per scientist.

We then ran our analysis on the sum of both types of memories ("Total Nb. memory events in academic publications" and "Total Nb. popular memory events") as well as a third category, the sum of both ("Total Nb. memory events"). For simplicity, we included only academic memories in the main body of the manuscript (Table 7). Here, however, we present identical specifications using popular and academic memories as outcomes (Tables D2 and D3, respectively).

The results suggest a broadly similar, but with attenuated magnitudes and noisier estimates for the main coefficients of interest. Similar to the results presented in Table 7, popular memories are also positively correlated with measures of eminence, namely publications, but the point estimates for the relationships are smaller and the standard errors are relatively larger, as seen in column 5. Similar to the academic memories in Table 7, membership in the NAS is far and away the best predictor of popular memories, as is evident in columns 2-7. Finally, self-promotion correlates with popular memories as it does with academic ones, but the point estimate is again slightly smaller. As total memories are the sum of popular and academic memories, the results naturally lie in between the two. Here, the relationship between publication count and memories becomes much stronger relative to popular memories, but not as strong as in the case of academic memories. Membership to the NAS remains large and significant throughout, as does self-promotion. Taken together, all three suggest a similar conclusion to that discussed in Section 4.4.1.

Table D1: Summary Statistics of Memory Events (N=676)

Type of Memory	Specific Memory	Mean	Median	Std. Dev.	Min.	Max.
Academic	Festschrift	0.073	0	0.294	0	2
	Memory event in an academic journal	1.830	1	2.443	0	19
	NAS memoir	0.117	0	0.321	0	1
	Symposium	0.030	0	0.170	0	1
	Total Nb. Memory Events in Academic Publications	**2.050**	**1**	**2.597**	**0**	**20**
Popular	New York Times obituary	0.338	0	0.477	0	2
	Other newspaper obituary	0.477	0	0.972	0	17
	University web post	0.495	0	0.685	0	4
	Wikipedia page	0.248	0	0.432	0	1
	Total Nb. Popular Memory Events	**2.129**	**2**	**2.858**	**0**	**55**
Total	**Total Nb. Memory Events**	**4.179**	**3**	**4.354**	**0**	**61**

Note: The sample size of 676 comprises the original scientists used in the main results data set.

Table D2: Estimating the Determinants of Memorialization – Popular Memories

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(cmltv. citations at death)		0.051			-0.064			
		(0.037)			(0.074)			
Ln(cmltv. publications at death)			0.151***		0.247**	0.231**	0.377**	0.354*
			(0.044)		(0.108)	(0.094)	(0.176)	(0.196)
Ln(cmltv. funding at death)				-0.008	-0.013			
				(0.010)	(0.010)			
Member of the NAS		0.646***	0.630***	0.716***	0.680***	0.701***	0.580***	0.614***
		(0.107)	(0.113)	(0.104)	(0.103)	(0.089)	(0.136)	(0.108)
Ln(Nb. of past trainees)						-0.120		-0.095
						(0.117)		(0.104)
Ln(Nb. of past coauthors [non-trainees])							-0.268	-0.184
							(0.198)	(0.152)
Self-Promoter								-0.156
								(0.124)
Female	0.071	0.115	0.164	0.069	0.150	0.148	0.218*	0.208*
	(0.127)	(0.112)	(0.112)	(0.114)	(0.113)	(0.110)	(0.124)	(0.121)
Death is Sudden	0.023	0.023	0.028	0.017	0.025	0.017	0.027	0.024
	(0.127)	(0.130)	(0.131)	(0.131)	(0.129)	(0.120)	(0.127)	(0.118)
Nb. of Scientists	647	647	647	647	647	647	647	647
Pseudo-R^2	0.056	0.098	0.101	0.097	0.104	0.105	0.108	0.115

Note: Estimates stem from Poisson specifications. The dependent variable is the total number of popular memories created for a scientist posthumously. A popular memory is a university web post, New York Times obituary, other newspaper obituary, Wikipedia page, or miscellaneous online obituary. All models include controls for degree type, death year, six age bins as well as the interaction terms for each age bin and sudden death. Self-Promoter is an indicator variable that corresponds to the top quartile of the distribution of unrelated citations as a percentage of self-citations (averaged over each deceased scientist's entire body of work) in our sample. Additionally, the specification in column (8) includes as a control the fraction of *related* self-citations as proportion of all self-citation (similarly averaged over the entire body of work of the focal scientist), to account for his/her propensity to cite his/her own work more generally. The sample size of 647 consists of the 676 original scientists less 29 scientists for which we were unable to confirm the cause of death (anticipated or sudden).

Robust standard errors in parentheses, clustered at the level of the star scientist. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D3: Estimating the Determinants of Memorialization – Total Memories

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(cmltv. citations at death)		0.154***			-0.030			
		(0.032)			(0.059)			
Ln(cmltv. publications at death)			0.324***		0.370***	0.385***	0.503***	0.471***
			(0.038)		(0.081)	(0.063)	(0.112)	(0.123)
Ln(cmltv. funding at death)				0.003	-0.009			
				(0.009)	(0.008)			
Member of the NAS		0.619***	0.640***	0.770***	0.664***	0.692***	0.604***	0.622***
		(0.081)	(0.082)	(0.079)	(0.079)	(0.074)	(0.093)	(0.088)
Ln(Nb. of past trainees)						-0.093		-0.072
						(0.075)		(0.070)
Ln(Nb. of past coauthors [non-trainees])							-0.213*	-0.139
							(0.124)	(0.099)
Self-Promoter								-0.097
								(0.115)
Female	-0.103	0.001	0.076	-0.074	0.068	0.066	0.112	0.103
	(0.115)	(0.098)	(0.095)	(0.101)	(0.095)	(0.095)	(0.098)	(0.099)
Death is Sudden	0.098	0.106	0.113	0.090	0.112	0.105	0.112	0.110
	(0.090)	(0.087)	(0.087)	(0.089)	(0.085)	(0.082)	(0.085)	(0.081)
Nb. of Scientists	647	647	647	647	647	647	647	647
Pseudo-R^2	0.053	0.140	0.155	0.124	0.157	0.158	0.161	0.167

Note: Estimates stem from Poisson specifications. The dependent variable is the total number of total memories created for a scientist posthumously. Total memories is the sum of both popular and academic memories. All models include controls for degree type, death year, six age bins as well as the interaction terms for each age bin and sudden death. Self-Promoter is an indicator variable that corresponds to the top quartile of the distribution of unrelated citations as a percentage of self-citations (averaged over each deceased scientist's entire body of work) in our sample. Additionally, the specification in column (8) includes as a control the fraction of *related* self-citations as proportion of all self-citation (similarly averaged over the entire body of work of the focal scientist), to account for his/her propensity to cite his/her own work more generally. The sample size of 647 consists of the 676 original scientists less 29 scientists for which we were unable to confirm the cause of death (anticipated or sudden).

Robust standard errors in parentheses, clustered at the level of the star scientist. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix E: List of 676 Deceased Scientists

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
Richard K. Gershon	1932 - 1983	MD	anticipated	Yale University	immunologic responses to tumor grafts
George Streisinger	1927 - 1984	PhD	sudden	University of Oregon	genetic mutations and the nervous system development in lower vertebrates
Arthur Cherkin	1913 - 1987	PhD	unknown	Sepulveda VA Medical Center	neurobiology of memory
Lucille S. Hurley	1922 - 1988	PhD	sudden	University of California — Davis	genetic and nutritional interactions in development
Toichiro Kuwabara	1920 - 1991	MD/PhD	sudden	Harvard Medical School	ophthalmology
Howard M. Temin	1934 - 1994	PhD	anticipated	University of Wisconsin	molecular biology and genetics of tumor viruses
Tsunao Saitoh	1949 - 1996	PhD	sudden	UCSD	altered protein kinases in alzheimer's disease
Paul C. MacDonald	1930 - 1997	MD	anticipated	University of Texas Southwestern Medical Center at Dallas	origin and interconversion of gonadal and adrenal streoid hormones
David J.L. Luck	1929 - 1998	MD/PhD	anticipated	Rockefeller University	microtubular systems in human cells
Robert H. Abeles	1926 - 2000	PhD	anticipated	Brandeis University	rational design of small-molecule inhibitors of enzymes
Keith Green	1940 - 2001	PhD	anticipated	Medical College of Georgia	ion and water movement in ocular tissues, ocular response to drugs
Charles A. Janeway, Jr.	1943 - 2003	MD	anticipated	Yale University	innate immunity and t lymphocyte biology
William H. Hildemann	1927 - 1983	PhD	anticipated	UCLA	mechanisms of immunoblocking versus tumor immunity
David T. Imagawa	1922 - 1991	PhD	sudden	Harbor-UCLA Medical Center	morphological conversion with leukemia viruses
Harold Koenig	1921 - 1992	MD/PhD	unknown	Northwestern University School of Medicine	neurology
Allastair M. Karmody	1937 - 1986	MD	anticipated	Albany Medical College	vascular surgery
Paul A. Obrist	1931 - 1987	PhD	anticipated	University of North Carolina at Chapel Hill	psychophysiology
William L. McGuire	1937 - 1992	MD	sudden	University of Texas HSC at San Antonio	mechanisms of hormonal control and growth and regression of mammary carcinoma
Edgar Haber	1932 - 1997	MD	anticipated	Harvard University School of Public Health	biological regulation of the renin-angiotensin system
Roger R. Williams	1944 - 1998	MD	sudden	University of Utah	genetics and epidemiology of coronary artery diseases
Lois K. Miller	1945 - 1999	PhD	anticipated	University of Georgia	genetics and molecular biology of baculoviruses
Frederick B. Bang	1916 - 1981	MD	sudden	Johns Hopkins School of Medicine	physician
Lewis W. Wannamaker	1923 - 1983	MD	sudden	University of Minnesota Medical School	clinical and epidemiologic aspects of streptococcal infections
Henry S. Kaplan	1918 - 1984	MD	anticipated	Stanford University School of Medicine	radiation-induced leukemia in the c57bl mouse
Joseph W. St. Geme, Jr.	1931 - 1986	MD	anticipated	University of Colorado HSC at Denver	pediatrics/infectious diseases
Richard P. Bunge	1932 - 1996	MD	anticipated	University of Miami	schwann cell biology and human spinal cord injury
Robert A. Good	1922 - 2003	MD/PhD	anticipated	University of South Florida College of Medicine	role of the thymus in immune system development
Thomas C. Chalmers	1917 - 1995	MD	anticipated	Mount Sinai Medical School	biostatistics
Barbara H. Bowman	1930 - 1996	PhD	unknown	University of Texas at Austin	biologist
Henry G. Kunkel	1916 - 1983	MD	sudden	Rockefeller University	identification of mhc class ii molecules
Nicholas R. DiLuzio	1926 - 1986	PhD	anticipated	Tulane University Medical School	physiology
Sandy C. Marks, Jr.	1937 - 2002	DDS/PhD	sudden	UMASS	bone cell biology
Morton I. Grossman	1919 - 1981	MD/PhD	anticipated	UCLA	studies on the etiology of peptic ulcer
Peter W. Lampert	1929 - 1986	MD	anticipated	UCSD	pathogenesis of virus-induced brain disease
James R. Neely	1936 - 1988	PhD	sudden	Penn State University	effects of diabetes and oxygen deficiency in regulation of metabolism in the heart
Frank J. Rauscher, Jr.	1931 - 1992	PhD	sudden	National Cancer Institute	cancer research
Carl M. Pearson	1919 - 1981	MD	anticipated	UCLA	studies in adjuvant-induced arthritis
Edward A. Smuckler	1931 - 1986	MD/PhD	anticipated	UCSF	cytochemical studies in liver injury
Catherine Cole-Beuglet	1936 - 1987	MD	anticipated	University of California — Irvine	ultrasonography of the breast
Edwin L. Bierman	1930 - 1995	MD	anticipated	University of Washington	metabolism of particulate fat in diabetes and atherosclerosis
G. Scott Giebink	1944 - 2003	MD	sudden	University of Minnesota	pathogenesis of otitis media and immunizations
Albert S. Gordon	1910 - 1992	PhD	sudden	New York University	internal medicine / hematology
Henryk M. Wisniewski	1931 - 1999	MD/PhD	sudden	SUNY Downstate Medical Center College of Medicine	pathogenesis of inflammatory demyelinating diseases
John S. O'Brien	1934 - 2001	MD	anticipated	UCSD	discovery of the gene responsible for tay-sachs disease
Kenneth L. Melmon	1934 - 2002	MD	sudden	Stanford University	autacoids as pharmacologic modifiers of immunity
Gregory Mooser	1942 - 2003	DDS/PhD	anticipated	University of Southern California	characterization of glucosyltranserase enzymes secreted by oral bacteria
Frederick Stohlman, Jr.	1929 - 1974	MD	sudden	Tufts University	medicine
Jerome R. Vinograd	1913 - 1976	PhD	sudden	California Institute of Technology	biochemistry and molecular biology
Michelangelo G.F. Fuortes	1917 - 1977	MD	sudden	Laboratory of Neurophysiology of the National Institute of Neurological Diseases and Stroke	neurology
Jerome T. Pearlman	1933 - 1979	MD	anticipated	UCLA	laboratory studies of retinal degenerations
David Pressman	1916 - 1980	PhD	sudden	State University of New York, Buffalo and Niagara University	immunochemistry
Richard C. Lillehei	1928 - 1981	MD/PhD	sudden	University of Minnesota	surgery
E. Jack Wylie	1918 - 1982	MD	sudden	UCSF	development of techniques for the treatment and management of chronic visceral ischemia
Roland L. Phillips	1937 - 1987	MD/PhD	sudden	Loma Linda University School of Medicine	role of lifestyle in cancer and cardiovascular disease among adventists
Jack Orloff	1921 - 1988	MD	anticipated	Division of Intramural Research at the National Heart, Lung, and Blook Institute	renal physiologist
Joaquim Puig-Antich	1944 - 1989	MD	sudden	University of Pittsburgh	psychobiology and treatment of child depression
Jean Mayer	1920 - 1993	PhD	sudden	Tufts University	nutritionist
Robert M. Joy	1941 - 1995	PhD	anticipated	University of California — Davis	pesticide induced changes in central nervous function
Paul M. Gallop	1927 - 1996	PhD	anticipated	Harvard Medical School/Children's Hospital	protein structure and collagen maturation
Richard Gorlin	1926 - 1997	MD	anticipated	Mount Sinai School of Medicine	studies of coronary blood flow and myocardial metabolism
Robert L. Summitt	1932 - 1998	MD	unknown	University of Tennessee, Memphis	pediatrics
Russell Ross	1929 - 1999	DDS/PhD	anticipated	University of Washington School of Medicine	response-to-injury origins of atherosclerosis
Donald J. Reis	1931 - 2000	MD	anticipated	Weill Medical College — Cornell University	neural control of blood circulation

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
Victor J. Ferrans	1937 - 2001	MD/PhD	sudden	NIH	myocardial and vascular pathobiology
Christopher A. Dawson	1942 - 2003	PhD	sudden	Medical College of Wisconsin	pulmonary hemodynamics
George G. Glenner	1927 - 1995	MD	anticipated	UCSD	molecular structure of the amyloid protein
Jiri Palek	1934 - 1998	MD	anticipated	Tufts University	membrane properties of abnormal red cells
Herman M. Kalckar	1908 - 1991	MD/PhD	sudden	Boston University	biochemistry
Frank A. Oski	1932 - 1996	MD	anticipated	Johns Hopkins	physician (blood disease and cancer)
John R. Williamson	1934 - 2000	PhD	anticipated	University of Pennsylvania School of Medicine	molecular mechanisms of hormonal signal transduction
Sydney E. Salmon	1936 - 1999	MD	anticipated	University of Arizona	quantitative method for evaluating changes in myeloma tumor mass
Alan P. Wolffe	1959 - 2001	PhD	sudden	NIH	role of dna methylation in regulating gene expression in normal and pathological states
Matthew I. Suffness	1942 - 1995	PhD	anticipated	National Cancer Institute	cancer research
C. Henry Kempe	1922 - 1984	MD	unknown	C. Henry Kempe National Center for the Prevention and Treatment of Child	child abuse and neglet
Milton Kern	1925 - 1987	PhD	anticipated	Abuse and Neglect;	immunology
Dante G. Scarpelli	1927 - 1998	MD/PhD	anticipated	Northwestern University School of Medicine	metabolism of pancreatic carcinogens
Gerald D. Aurbach	1927 - 1991	MD	sudden	NIH	bone metabolism and calcium homeostasis
Werner H. Kirsten	1925 - 1992	MD	sudden	NCI-Frederick Cancer Research and Development Center in Frederick, Maryland	cancer research
Allan C. Wilson	1934 - 1991	PhD	anticipated	University of California — Berkeley	use of molecular approaches to understand evolutionary change
Irving J. Selikoff	1915 - 1992	MD	anticipated	Mount Sinai School of Medicine	asbestos and cancer
Arnold F. Brodie	1923 - 1981	PhD	unknown	University of Southern California	microbiology
Takis S. Papas	1935 - 1999	PhD	sudden	Medical University of South Carolina	characterization of ets genes and retroviral onc genes
Harland G. Wood	1907 - 1991	PhD	anticipated	Case Western Reserve University School of Medicine	heterotrophic carbon dioxide fixation
George B. Craig, Jr.	1930 - 1995	PhD	sudden	University of Notre Dame	genetics and reproductive biology of aedes mosquitoes
John P. Merrill	1917 - 1984	MD	sudden	Harvard Medical School/Brigham & Women's Hospital	role of the immune system in kidney transplantation
Sheldon M. Wolff	1930 - 1994	MD	anticipated	Tufts University School of Medicine	treatment of fevers from infectious diseases like wegener's granulomatosis
Kenneth M. Moser	1929 - 1997	MD	anticipated	UCSD School of Medicine	clinical outcomes after pulmonary thromboendarterectomy
Thomas P. Dousa	1937 - 2000	MD/PhD	sudden	Mayo Clinic	cellular action of vasopressin in the kidney
Jeffrey M. Isner	1947 - 2001	MD	sudden	Tufts University	therapeutic angiogenesis in vascular medicine, cardiovascular laser phototherapy
Jon I. Isenberg	1937 - 2003	MD	anticipated	UCSD	duodenal mucosal bicarbonate secretion in human
Richard P. Nordan	1949 - 1998	PhD	sudden	NIH	immunologist and molecular biologist
Trudy L. Bush	1949 - 2001	PhD	sudden	University of Maryland School of Medicine	postmenopausal estrogen/progestins interventions
James R. Klinenberg	1934 - 1999	MD	sudden	UCLA	pathophysiology of gout and hyperuricemia
Matthew L. Thomas	1953 - 1999	PhD	sudden	Washington University in St. Louis	function and regulation of leukocyte surface glycoproteins
Alfred P. Wolf	1923 - 1998	PhD	anticipated	Brookhaven National Laboratory	synthesis of simple molecules in pure form and high specific activity for pet
Ronald G. Thurman	1941 - 2001	PhD	sudden	University of North Carolina	hepatic metabolism, alcoholic liver injury and toxicology
Emil T. Kaiser	1938 - 1988	PhD	sudden	Rockefeller University	mechanism of carboxypeptidase action
Hymie L. Nossel	1930 - 1983	MD/PhD	sudden	Columbia University	causes of thrombosis and the nature of hemostasis
Marian W. Fischman	1939 - 2001	PhD	anticipated	Columbia University	behavioral pharmacology of cocaine
Charles D. Heidelberger	1920 - 1983	PhD	anticipated	University of Southern California	effects of fluorinated pyrimidines on tumors
Choh Hao Li	1913 - 1987	PhD	anticipated	UCSF School of Medicine	isolation and synthesis the human pituitary growth hormone
DeWitt S. Goodman	1930 - 1991	MD	sudden	Columbia University	lipid metabolism and its role in the development of heart and artery disease
Alvito P. Alvares	1935 - 2001	PhD	sudden	University of Bethesda, Maryland	pharmacology
Charlotte Friend	1921 - 1987	PhD	anticipated	Mount Sinai School of Medicine	tissue studies of murine virus-induced leukemia
Gerald L. Klerman	1928 - 1992	MD	anticipated	Weill Medical College — Cornell University	phsychological studies of depression, schizophrenia and panic and other anxiety disorders
J. David Robertson	1922 - 1995	MD/PhD	anticipated	Duke University School of Medicine	electron microscopy of cell membranes
Fred H. Allen, Jr.	1912 - 1987	MD	sudden	New York Blood Center	blood grouping
Markku Linnoila	1947 - 1998	MD/PhD	anticipated	NIH	studies on the biological bases of impulsivity and aggression
Muriel R. Steele	1930 - 1979	MD	anticipated	University of California- San Francisco	surgery
J. Weldon Bellville	1926 - 1983	MD	anticipated	UCLA	dynamic isolation studies of control of respiration
Kwan C. Tsou	1922 - 1985	PhD	sudden	University of Pennsylvania	chemistry and pharmacology
Nathan O. Kaplan	1917 - 1986	PhD	sudden	UCSD	isolation and structure determination of coenzyme a
Amico Bignami	1930 - 1994	MD	anticipated	Harvard Medical School	brain specific protein in astrocytes
Christian B. Anfinsen, Jr.	1916 - 1995	PhD	sudden	Johns Hopkins University	protein structure and protein folding
Hans J. Müller-Eberhard	1927 - 1998	MD	anticipated	Scripps Research Institute	identification of proteins and reaction mechanisms of the complement system
Louis V. Avioli	1931 - 1999	MD	anticipated	Washington University School of Medicine	mineral and skeletal metabolism in diabetes, kidney, and gastrointestinal disorders
Edward Herbert	1926 - 1987	PhD	anticipated	Oregon Health & Science University	regulation of expression of opioid peptides and receptors
Efraim Racker	1913 - 1991	MD	sudden	Cornell University	identifying and purifying factor 1, the first part of the atp synthase enzyme
Simon J. Pilkis	1942 - 1995	MD/PhD	sudden	University of Minnesota	carbohydrate metabolism and diabetes
Thomas W. Smith	1936 - 1997	MD	anticipated	Harvard Medical School/Brigham & Women's Hospital	mechanism and reversal studies of digitalis
Harriet P. Dustan	1920 - 1999	MD	anticipated	University of Vermont	hypertension specialist
Don C. Wiley	1944 - 2001	PhD	sudden	Harvard University	viral membrane and glycoprotein structure
Eva U.J. Paucha	1949 - 1988	PhD	anticipated	Dana Farber Cancer Institute	mechanism of transformation by sv40 large t antigen
Gareth M. Green	1931 - 1998	MD/PhD	anticipated	Harvard University School of Public Health	role of alveolar macrophages in pulmonary defense mechanisms
Geoffrey H. Bourne	1909 - 1988	PhD	sudden	St. George's University School of Medicine in Grenada	nutrition and primates
Edward W. Moore	1930 - 1999	MD	anticipated	Medical College of Virginia	pathophysiology of the billiary tract and gallbladder
Ernst A. Noltmann	1931 - 1986	MD	anticipated	University of California — Riverside	biochemical and physical characterization of phosphoglucose isomerase
Robert M. Pratt, Jr.	1942 - 1987	PhD	sudden	NIEHS / University of North Carolina at Chapel Hill	biochemistry
Hans Popper	1903 - 1988	MD/PhD	anticipated	Mount Sinai School of Medicine	correlation of structure and function in liver disease
Thomas P. Hackett, Jr.	1928 - 1988	MD	sudden	Massachusetts General Hospital / Harvard Medical School	psychosomatic medicine
Carl W. Gottschalk	1922 - 1997	MD	sudden	University of North Carolina at Chapel Hill School of Medicine	micropuncture studies of mammallian renal system
William L. Chick	1938 - 1998	MD	anticipated	UMASS	studies of islet and beta cells in pancreatic transplantation

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
Elizabeth Stern	1915 - 1980	MD	anticipated	UCLA	effects of steroid contraception on the ovary
Harold Edelhoch	1922 - 1986	PhD	anticipated	National Institute of Arthristis, Diabetes, and Digestive Kidnesy Diseases	oncology and biochemistry
Ephraim Donoso	1917 - 1988	MD	unknown	Mount Sinai School of Medicine	cardiology
Norton B. Gilula	1944 - 2000	PhD	anticipated	Scripps Research Institute	cell junction biosynthesis and biogenesis/cell-cell communication
George Khoury	1943 - 1987	MD	anticipated	NIH	genetics of simian virus 40, human papovavirus and hiv
Charles G. Moertel	1927 - 1994	MD	anticipated	Mayo Clinic	clinical treatments of gastrointestinal cancer
Bruce W. Erickson	1942 - 1998	PhD	anticipated	University of North Carolina at Chapel Hill	engineering of nongenetic beta proteins
Robert J. Fass	1939 - 2002	MD	anticipated	Ohio State University	in vitro methods to test antimicrobial susceptibility of infectious agents
Ahmad I. Bukhari	1943 - 1983	PhD	sudden	Cold Spring Harbor Laboratory	life cycle of mutator phage μ
B. Frank Polk	1942 - 1988	MD	anticipated	Johns Hopkins University	epidemiology of hiv infection
Frank Lilly	1930 - 1995	PhD	anticipated	Albert Einstein College of Medicine	role of hereditary factors in governing susceptibility to cancer-causing agents
Donald S. Fredrickson	1924 - 2002	MD	sudden	National Library of Medicine	physician
Eva J. Neer	1937 - 2000	MD	anticipated	Harvard Medical School/Brigham & Women's Hospital	regulation and cellular levels of g protein subunits
Sol Spiegelman	1914 - 1983	PhD	anticipated	Columbia University College of Physicians and Surgeons	nucleic acid hybridization
Nina S. Braunwald	1928 - 1992	MD	anticipated	Harvard Medical School/Brigham & Women's Hospital	development of prosthetic heart valves for children
Nava Sarver	1951 - 2001	PhD	anticipated	National Institute of Health	aids
Jeffrey M. Hoeg	1952 - 1998	MD	sudden	NHLBI's Molecular Disease Branch	cell biology
Joseph E. Coleman	1930 - 1999	MD/PhD	anticipated	Yale University	structure and function of metalloenzyme synthesis
Paul B. Sigler	1934 - 2000	MD/PhD	sudden	Yale University	structural analysis of biological macromolecules
Gerald L. Stoner	1943 - 2002	PhD	sudden	NINDS	neurotoxicology
Murray Rabinowitz	1927 - 1983	MD	anticipated	University of Chicago	mitochondrial assembly and replication
Samuel Sutton	1921 - 1986	PhD	sudden	University of Chicago	biometrics
Murray B. Bornstein	1918 - 1995	MD	sudden	Albert Einstein College of Medicine of Yeshiva University	copolymer as a protective treatment for the exacerbation of multiple sclerosis
Harold C. Neu	1934 - 1998	MD	anticipated	Columbia University	surface enzymes in bacteria
J. Murray Steele	1900 - 1969	MD	sudden	New York University	medicine
Harvey M. Patt	1918 - 1982	PhD	anticipated	UCSF	ultra-high dose rates in experimental radiotherapy
John J. Pisano	1929 - 1985	PhD	sudden	Laboratory of Chemistry, National Heart, Lung, and Blood Institute	physiological chemistry
Edward H. Kass	1917 - 1990	MD/PhD	anticipated	Harvard Medical School/Brigham & Women's Hospital	mechanism of toxic shock syndrome
Fred I. Gilbert, Jr.	1920 - 1995	MD	unknown	University of Hawaii	clinical studied of hyper- and hypothyroidism
James K. McDougall	1931 - 2003	PhD	anticipated	University of Washington/FHCRC	role of dna viruses in cancer
Roger O. Eckert	1934 - 1986	PhD	anticipated	UCLA	ionic and metabolic mechanisms in neuronal excitability
James W. Maas	1929 - 1995	MD	unknown	University of Texas Health Sciences Center at San Antonio	psychiatrist
Miriam M. Salpeter	1929 - 2000	PhD	anticipated	Cornell University	neurobiology of myasthenia gravis
Mary Lou Clements	1946 - 1998	MD	sudden	Johns Hopkins University	development of aids vaccines
Donald B. Hackel	1921 - 1994	MD	anticipated	Duke University	cardiac pathology
Vincent Massey	1926 - 2002	PhD	sudden	University of Michigan Medical School	biological oxidation mechanisms of proteins that contain riboflavin
Gordon Guroff	1933 - 1999	PhD	sudden	NICHD (National Institute of Child Health and Human Development)	biochemistry
Merton Bernfield	1938 - 2002	MD	anticipated	Harvard Medical School/Children's Hospital	nature and interactions of cell surface proteoglycans during morphogenesis
Werner Henle	1910 - 1987	MD	anticipated	University of Pennsylvania School of Medicine	serologic response to epstein-barr virus infection
Aaron Janoff	1930 - 1988	PhD	anticipated	SUNY HSC at Stony Brook	pathology of smoking and emphysema
Joachim G. Liehr	1942 - 2003	PhD	anticipated	University of Texas Medical Branch at Galveston	mechanism of estrogen-induced carcinogenesis
Mortimer B. Lipsett	1921 - 1985	MD	anticipated	NIH	steroid metabolic conversions in human subjects
John H. Walsh	1938 - 2000	MD	sudden	UCLA	gastrointestinal hormones, gastric acid production and peptic ulcer disease
Erhard Gross	1928 - 1981	PhD	sudden	National Health Institute	chemistry
C. Clark Cockerham	1921 - 1996	PhD	unknown	North Carolina State University	statistical genetics
Lubomir S. Hnilica	1929 - 1986	PhD	sudden	Vanderbilt University	nuclear antigens in human colorectal cancer
Julius Marmur	1926 - 1996	PhD	anticipated	Albert Einstein College of Medicine of Yeshiva University	genetics and biochemistry of cellular regulation
Ronald S. Wilson	1933 - 1986	PhD	sudden	University of Louisville	behavioral genetics
Edward W. Hook, Jr.	1924 - 1998	MD	sudden	University of Virginia	physician
Paul P. Carbone	1931 - 2002	MD	sudden	University of Wisconsin Medical School	treatment and prevention of hodgkin's disease and early breast cancer
Hamish N. Munro	1915 - 1994	MD/PhD	anticipated	Tufts University	nutrition scientist
Sidney Leskowitz	1923 - 1991	PhD	anticipated	Tufts Medical School	immunology
Roland D. Ciaranello	1943 - 1994	MD	sudden	Stanford University	molecular neurobiology and developmental disorders
Daniel Mazia	1912 - 1996	PhD	anticipated	Stanford University	isolation of the mitotic apparatus
Julian M. Davidson	1931 - 2001	PhD	anticipated	Stanford University	physiological bases of human sexuality
Sidney H. Ingbar	1925 - 1988	MD	anticipated	Harvard Medical School/Beth Israel Medical Center	physiology of the thyroid gland and its clinical diseases
Michael Solursh	1942 - 1994	PhD	anticipated	University of Iowa School of Medicine	extracellular matrix and cell migration
Herbert J. Rapp	1923 - 1981	PhD	sudden	National Cancer Institute	immunologist and cancer research
Paul A. Srere	1925 - 1999	PhD	sudden	University of Texas Southwestern Medical Center at Dallas	cell metabolism and the krebs tca cycle
Frank L. Horsfall, Jr.	1906 - 1971	MD	anticipated	Sloan-Kettering Institute	clinician and virologist
J. Werner Braun	1914 - 1972	PhD	unknown	Rutgers University	microbiology
Elijah Adams	1918 - 1979	MD	unknown	University of Maryland, Baltimore	biochemistry
Albert Dorfman	1916 - 1982	MD/PhD	anticipated	University of Chicago	biochemistry of connective tissues
Abraham I. Braude	1917 - 1984	MD/PhD	sudden	UCSD	pathogenesis and treatment of life-threatening septic shock
Albert L. Lehninger	1917 - 1986	PhD	anticipated	Johns Hopkins University School of Medicine	structure and function of mitochondria
Michael A. Kirschenbaum	1944 - 1997	MD	anticipated	UC Irvine	prostaglandins and kidney medicine
George F. Solomon	1931 - 2001	MD	sudden	UCLA	psychiatry and biobehavioural sciences
Joseph W. Beard	1906 - 1983	MD	unknown	Duke University	biology, animal viruses, cancer induced by viruses
Ralph R. Cavalieri	1932 - 2001	MD	sudden	UCSF	utilization of tyrosine by the thyroid gland

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
Philip Handler	1917 - 1981	PhD	anticipated	Duke University	biochemist
N. Raphael Shulman	1925 - 1996	MD	anticipated	NIDDK (NIH)	hematology (biochemistry)
Dale E. McFarlin	1936 - 1992	MD	sudden	NIH	neuroimmunological studies of multiple sclerosis
James M. Felts	1923 - 1988	PhD	sudden	UCSF	synthesis and processing of plasma lipoproteins
John P. Merlie	1945 - 1995	PhD	sudden	Washington University in St. Louis	molecular genetics of the acetylcholine receptor
Wallace P. Rowe	1926 - 1983	MD	anticipated	NIH	genetic basis of disease in murine leukemia viruses
Nelson Butters	1937 - 1995	PhD	anticipated	UCSD	cognitive deficits related to chronic alcoholism
John B. Penney, Jr.	1947 - 1999	MD	sudden	Harvard Medical School/MGH	receptor mechanisms in movement disorder pathophysiology
John C. Seidel	1933 - 1988	PhD	sudden	Boston Biomedical Research Institute	actin-myosin interaction in pulmonary smooth muscle
William H. Oldendorf	1925 - 1992	MD	sudden	UCLA	x-ray shadow radiography and cerebral angiography
Kenneth W. Sell	1931 - 1996	MD/PhD	anticipated	Emory University	immunology
Richard J. Winzler	1916 - 1972	PhD	sudden	Florida State University	biochemistry
Andrew C. Peacock	1921 - 1985	PhD	anticipated	National Cancer Institute	molecules, dna method
Harold A. Menkes	1938 - 1987	MD	sudden	Johns Hopkins University	occupational and environmental lung disease
Norman Kretchmer	1923 - 1995	MD/PhD	anticipated	University of California, Berkeley	obstetrics, pediatrics
John C. Liebeskind	1935 - 1997	PhD	anticipated	UCLA	behavioral and electrophysiological studies of pain
Thomas B. Fitzpatrick	1919 - 2003	MD/PhD	anticipated	Harvard Medical School/MGH	dynamics of epidermal pigmentation
Menek Goldstein	1924 - 1997	PhD	sudden	New York University School of Medicine	purification of enzymes in the catecholamine synthetic pathway
Paul J. Scheuer	1915 - 2003	PhD	anticipated	University of Hawaii	organic chemistry
Loretta L. Leive	1936 - 1986	PhD	anticipated	NIH	microbiology
George Némethy	1934 - 1994	PhD	anticipated	Mount Sinai School of Medicine	methods to analyze and predict the structures of protein molecules
Lonnie D. Russell, Jr.	1944 - 2001	PhD	sudden	Southern Illinois University School of Medicine	filament regulation of spermatogenesis
Samuel A. Latt	1938 - 1988	MD/PhD	sudden	Harvard Medical School/Children's Hospital	genetic and cytogenetic studies of mental retardation
Carl Monder	1928 - 1995	PhD	sudden	Population Council	corticosteroid metabolism in juvenile hypertension
Gary J. Miller	1950 - 2001	MD/PhD	sudden	University of Colorado HSC	vitamin d receptors in the growth regulation of prostate cancer cells
Takeo Kakunaga	1937 - 1988	PhD	sudden	National Cancer Institute	genesis of human cancer
Richard E. Heikkila	1942 - 1991	PhD	sudden	UMDNJ Robert Wood Johnson Medical School	oxidation-reduction reactions and the dopamine receptor system
Roderich Walter	1937 - 1979	PhD	anticipated	University of Illinois	physiology and biophysics
Daniel Rudman	1927 - 1994	MD	sudden	Medical College of Wisconsin	adipokinetic substances of the pituitary gland
J. Calvin Giddings	1930 - 1996	PhD	anticipated	University of Utah, Salt Lake City	chemistry
Ethan R. Nadel	1941 - 1998	PhD	anticipated	Yale University	thermoregulation during exercise and heat exposure
Bertram Sacktor	1922 - 1988	PhD	sudden	National Institute on Aging in Baltimore	gerontology
Gerald T. Babcock	1946 - 2000	PhD	anticipated	Michigan State University	bioenergetic mechanisms in multicenter enzymes
George Winokur	1925 - 1996	MD	anticipated	University of Iowa College of Medicine	genetics of bipolar disease, mania, alcoholism and other psychiatric diseases
Marian E. Koshland	1921 - 1997	PhD	anticipated	University of California — Berkeley	biochemical methods to examine the immune response
John J. Jeffrey, Jr.	1937 - 2001	PhD	sudden	Albany Medical College	mechanism of action and the physiologic regulation of mammalian collagenases
Richard N. Lolley	1933 - 2000	PhD	sudden	USC (Keck School of Medicine)	neurochemistry
Henry R. Mahler	1921 - 1983	PhD	anticipated	Indiana University	chemistry
Michael J. Goldstein	1930 - 1997	PhD	anticipated	UCLA	contributing factors to the onset of schizophrenia
Ernest Borek	1911 - 1986	PhD	unknown	City University of New York	chemistry
Seymour Fisher	1922 - 1996	PhD	sudden	State University of New York at Syracuse	psychology
Gordon M. Tomkins	1926 - 1975	PhD	anticipated	University of California at San Francisco	hormoe research
Jordi Folch-Pi	1911 - 1979	MD	sudden	Massachusetts General Hospital	neurochemistry
W. Dean Warren	1924 - 1989	MD	anticipated	Emory University	surgery
Roger T. Kelleher	1926 - 1994	PhD	unknown	Harvard Medical School	pharmacology
Hugh L. Keegan	1916 - 1980	PhD	anticipated	University of Mississippi Medical Center	preventive medicine
Andrew G. Morrow	1923 - 1982	MD	unknown	National Heart, Lung and Blood Institute	surgery
Teruzo Konishi	1920 - 1984	MD/PhD	anticipated	NIEHS	physiological and biophysical functions of the inner ear
C. Richard Taylor	1939 - 1995	PhD	anticipated	Harvard University	energetics of animal locomotion
Roy H. Steinberg	1935 - 1997	MD/PhD	anticipated	UCSF	pigment epithelium interactions with neural retina
Fredric S. Fay	1943 - 1997	PhD	sudden	UMASS	generation and regulation of force in smooth muscle
Verne M. Chapman	1938 - 1995	PhD	sudden	Roswell Park Cancer Institute/SUNY Buffalo	development of cumulative multilocus map of mouse chromosomes
Priscilla A. Campbell	1940 - 1998	PhD	anticipated	University of Colorado HSC/Nat. Jewish center	cell biology of the immune response to bacteria
Sol Levine	1922 - 1996	PhD	sudden	New England Medical Center	medical sociology
Donald A. Pious	1930 - 1998	MD	anticipated	University of Washington School of Medicine	somatic cell genetic analysis of human immune response genes
Elizabeth A. Rich	1952 - 1998	MD	sudden	Case Western Reserve University School of Medicine	natural history of lymphocytic alveolitis in hiv disease
Peter Safar	1924 - 2003	MD	anticipated	University of Pittsburgh	clinical studies of brain resuscitation
Belding H. Scribner	1921 - 2003	MD	sudden	University of Washington	dialysis in the treatment of chronic uremia
S. Morris Kupchan	1922 - 1976	PhD	unknown	University of Virginia	chemistry
Edward J. Sachar	1933 - 1984	MD	anticipated	Columbia University	chemicals in mental illness
Bernard N. Fields	1938 - 1995	MD	anticipated	Harvard Medical School/Brigham & Women's Hospital	genetic and molecular basis of viral injury to the nervous system
G. Jeanette Thorbecke	1929 - 2001	MD/PhD	sudden	New York University School of Medicine	histologic and functional aspects of lymphoid tissue development
Fritz A. Lipmann	1899 - 1986	MD/PhD	anticipated	Rockefeller University	biochemistry
Theodore S. Zimmerman	1937 - 1988	MD	anticipated	Scripps Research Institute	platelet/plasma protein interaction in blood coagulation
David G. Marsh	1940 - 1998	PhD	anticipated	Johns Hopkins University	genetics of allergy and asthma
A. Arthur Gottlieb	1937 - 1998	MD	sudden	Tulane University School of Medicine	role of macrophage nucleic acid in antibody production
Joseph B. Warshaw	1936 - 2003	MD	anticipated	University of Vermont College of Medicine	developmental neurobiology of respiratory control
Frederic C. Bartter	1914 - 1983	MD	sudden	University of Texas HSC at San Antonio	interaction between the kidney and various endocrine systems
Edward C. Franklin	1928 - 1982	MD	anticipated	New York University	structure and properties of rheumatoid antibodies

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
James A. Campbell	1917 - 1983	MD	sudden	Rush-Presbyterian-St Luke's Medical Center	internal medicine and cardiology
Robert J. Stoller	1924 - 1991	MD	sudden	UCLA	clinical studies of gender identity
Thomas F. Burks, II	1938 - 2001	PhD	sudden	University of Texas HSC at Houston	central and peripheral neuropeptide pharmacology
Dolph O. Adams	1939 - 1996	MD/PhD	sudden	Duke University	development and regulation of macrophage activation
Mette Strand	1937 - 1997	PhD	anticipated	Johns Hopkins University	parasite immunochemistry and vaccine development
Joel D. Meyers	1944 - 1991	MD	anticipated	University of Washington/FHCRC	infections caused by suppression of the immune system in organ transplant and aids patients
Thomas D. Kinney	1909 - 1977	MD	anticipated	Duke University	pathology
Bernard D. Davis	1916 - 1994	MD	anticipated	Harvard Medical School	microbiology
F. Brantley Scott, Jr.	1930 - 1991	MD	sudden	Baylor University College of Medicine	urologist
Fred Karush	1913 - 1994	PhD	anticipated	University of Pennsylvania Medical School	microbiology
Thomas Francis, Jr.	1900 - 1969	MD	sudden	University of Michigan	physician, virologist, and epidemiologist
William J. Mellman	1928 - 1980	MD	anticipated	University of Pennsylvania	human genetics and pediatrics
Charles W. Todd	1918 - 1987	PhD	anticipated	City of Hope Medical Center	immunology / biochemistry
Tsoo E. King	1917 - 1990	PhD	unknown	University of Pennsylvania School of Medicine	bioenergetic apparatus in heart mitochondria
Jonas E. Salk	1914 - 1995	MD	sudden	Salk Inst. Biol. Studies	effective vaccine for polio
Gilda H. Loew	1931 - 2001	PhD	anticipated	Molecular Research Institute	computational investigation of the structural and functional aspects of heme proteins and enzymes
Lloyd J. Filer, Jr.	1919 - 1997	MD/PhD	sudden	University of Iowa	pediatrics
Neil S. Jacobson	1949 - 1999	PhD	sudden	University of Washington	marital therapy, domestic violence, and the treatment of depression
Philip J. Fialkow	1934 - 1996	MD	sudden	University of Washington	origins of myeloid leukemia tumors
Marian W. Kies	1915 - 1988	PhD	sudden	NIH	biochemistry
Pierre M. Galletti	1927 - 1996	MD/PhD	sudden	Brown Medical School	synthesis of artificial lung and kidney systems
Mu-En Lee	1954 - 2000	MD/PhD	sudden	Harvard Medical School/MGH	characterization of vascular smooth muscle lim protein
Ira Herskowitz	1946 - 2003	PhD	anticipated	UCSF	genetics of yeast mating type
Robert M. Macnab	1940 - 2003	PhD	sudden	Yale University	sequence analysis and function of bacterial flagellar motor
C. Andrew L. Bassett	1924 - 1994	MD/PhD	anticipated	Bioelectric Research Center / Columbia University	orthopedic surgery
Ernest Witebsky	1901 - 1969	MD	sudden	University of Buffalo	immunology
Lyman C. Craig	1906 - 1974	PhD	unknown	The Rockefeller Institute (NY)	biochemistry
George C. Cotzias	1918 - 1977	MD	anticipated	Cornell University Medical College	neurology
Robert S. Krooth	1929 - 1980	PhD	sudden	College of Physicians and Surgeons of Columbia University	human genetics and development
Thomas R. Johns, 2nd	1924 - 1988	MD	sudden	University of Virginia School of Medicine	physiological studies of myasthenia gravis
Marshall H. Becker	1940 - 1993	PhD	anticipated	University of Michigan, Ann Arbor	elaboration of the health belief model
Samuel W. Perry, 3rd	1941 - 1994	MD	anticipated	Weill Medical College — Cornell University	psychological course of prolonged infection among aids patients
Howard S. Tager	1945 - 1994	PhD	sudden	University of Chicago	biochemical structure, action, regulation and degradation of the insulin and ghcagon molecules
Andrew P. Somlyo	1930 - 2003	MD	sudden	University of Virginia School of Medicine	vasomotor function of smooth muscle and their relation to heart disease
Walter F. Heiligenberg	1938 - 1994	PhD	sudden	UCSD	neuroethological studies of electrolocation
Donald C. Shreffler	1933 - 1994	PhD	sudden	Washington University in St. Louis	organization and functions of h-2 gene complex
Gisela Mosig	1930 - 2003	PhD	anticipated	Vanderbilt University	genetics
Robert F. Spencer	1949 - 2001	PhD	anticipated	Medical College of Virginia	neuroanatomy of the oculomotor system
Harry A. Waisman	1912 - 1971	MD/PhD	sudden	University of Wisconsin	pediatrics and child development
Raymond T. Carhart	1912 - 1975	PhD	sudden	Northwestern University	audiology and otolaryngology
Frederick H. Carpenter	1918 - 1982	PhD	anticipated	University of California — Berkeley	mechanism of leucine aminopeptidase
Peter D. Klein	1927 - 2001	PhD	unknown	Baylor College of Medicine	pediatrics/gastroenterology
Oscar A. Kletzky	1936 - 1994	MD	anticipated	UCLA	ameliorating effects of estrogen replacement therapy on cerebral blood flow and sleep
Myron L. Bender	1924 - 1988	PhD	unknown	Northwestern University	chemist
James N. Gilliam	1936 - 1984	MD	anticipated	University of Texas Southwestern Medical Center at Dallas	cutaneous lupus erythematosus pathogenesis mechanisms
Bernard Sass	1935 - 1989	MD	anticipated	National Cancer Institute	veterinary pathologist
Ernst Freese	1925 - 1990	PhD	sudden	National Institute of Neurological Disorders and Stroke (NIH in Bethesda)	molecular biology
David M. Maurice	1922 - 2002	PhD	anticipated	Columbia University College of Physicians and Surgeons	interference theory of corneal transparency
Eugene P. Cronkite	1914 - 2001	MD	anticipated	Brookhaven National Laboratory	hematopoiesis and radiation injury
Gerald P. Rodnan	1927 - 1983	MD	sudden	University of Pittsburgh	renal transport if uric acid and protein
Norman P. Salzman	1926 - 1997	PhD	anticipated	NIH	virologist
Ruth Sager	1916 - 1997	PhD	anticipated	Harvard Medical School/DFCI	role of tumor suppressor genes in breast cancer
Dennis Slone	1930 - 1982	MD	anticipated	Boston University Medical Center	epidemiology
Ramzi S. Cotran	1932 - 2000	MD	anticipated	Harvard Medical School/Brigham & Women's Hospital	mechanisms of immune, infectious, and vascular renal injury
Maurice Lev	1908 - 1994	MD	unknown	Rush Medical College, Chicago	pathology and cardiology
Joseph H. Ogura	1915 - 1983	MD	sudden	Washington University (St. Louis)	otolaryngology
Gerald Cohen	1930 - 2001	PhD	anticipated	Mount Sinai School of Medicine	neurology
Chaviva Isersky	1937 - 1986	PhD	anticipated	National Institute of Arthristis, Diabetes, and Digestive Kidnesy Diseases	arthritus and rheumatism
Henry Rapoport	1918 - 2002	PhD	sudden	University of California — Berkeley	total synthesis of heterocyclic drugs
Wallace H. Clark, Jr.	1924 - 1997	MD	sudden	Harvard Medical School	pathologist, dermatologist
David Spiro	1921 - 1974	MD/PhD	sudden	New York Medical College	pathology
Reidar F.A. Sognnaes	1911 - 1984	PhD	sudden	UCLA School of Dentistry	forensic scientist on dental records
Adolph I. Cohen	1924 - 1996	PhD	anticipated	Washington University in St. Louis	ophthalmology and anatomy
Clarence J. Gibbs, Jr.	1924 - 2001	PhD	sudden	NINDS	infectuous diseases of the nervous system
Carl C. Levy	1928 - 1981	PhD	anticipated	Baltimore Cancer Research Program (National Cancer Institute)	molecular biology
Helene S. Smith	1941 - 1997	PhD	anticipated	UCSF	malignant progression of the human breast/predictors of breast cancer prognosis
D. Eugene Strandness, Jr.	1928 - 2002	MD	sudden	University of Washington School of Medicine	ultrasonic duplex scanner for noninvasive vascular disease diagnosis
William S. Beck	1924 - 2003	MD	anticipated	Harvard Medical School	biochemistry of blood cell formation
David Tapper	1945 - 2002	MD	anticipated	University of Washington School of Medicine, Seattle, Washington	pediatric surgeon

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
Jonathan M. Mann	1943 - 1998	MD	sudden	Harvard University School of Public Health	aids prevention
Bernard R. Baker	1915 - 1971	PhD	sudden	UCSB	chemistry
Leonard R. Axelrod	1927 - 1975	PhD	unknown	Environmental Protection Agency	biochemistry
Cyril S. Stulberg	1919 - 1977	PhD	anticipated	Wayne State University School of Medicine	immunology and microbiology
Herbert F. Hasenclever	1924 - 1978	PhD	anticipated	National Institute of Allergy and Infectious Diseases / NIH	mycologist
William H. Tooley	1925 - 1992	MD	anticipated	UCSF School of Medicine	prevention and treatment of respiratory distress in neonates
J. Christian Gillin	1938 - 2003	MD	anticipated	UCSD	serotenergic mechanisms in sleep and depression
Dan H. Campbell	1907 - 1974	PhD	sudden	California Institute of Technology	immunochemistry
Melvin L. Marcus	1940 - 1989	MD	anticipated	UMASS	cardiology, heart disease, coronary vascular adaptations to myocardial hypertrophy
Charles E. Putman	1941 - 1999	MD	sudden	Duke University	radiologist
Kenneth J.W. Taylor	1939 - 2003	MD/PhD	unknown	Yale Medical School	diagnostic ultrasound imaging
Edgar E. Ribi	1920 - 1986	PhD	sudden	NIAID Lab in Hamilton, Montana	chemistry and biophysics
Alton Meister	1922 - 1995	MD	anticipated	Weill Medical College — Cornell University	amino acid and glutathione biochemistry
John A. Kirkpatrick, Jr.	1926 - 1994	MD	unknown	Harvard Medical School	radiologist
Peter M. Steinert	1945 - 2003	PhD	sudden	NIH	structures and interactions of the proteins characteristic of epithelial cells
Milton Orkin	1929 - 1999	MD	anticipated	University of Minnesota	dermatology
Lawrence Bogorad	1921 - 2003	PhD	sudden	Harvard University	biologist
Joseph F. Foster	1918 - 1975	PhD	sudden	Purdue University	biochemistry
James Olds	1922 - 1976	PhD	sudden	California Institute of Technology	psychology
Carl L. Larson	1909 - 1978	MD	unknown	University of Montana at Missoula	microbiology
Russell J. Barrnett	1920 - 1989	MD	sudden	Yale Medical School	cell biology
Brigitte A. Prusoff	1926 - 1991	PhD	unknown	Yale University	chemist
Nemat O. Borhani	1926 - 1996	MD	anticipated	University of Nevada at Reno	multicenter clinical studies of hypertension and cardiovascular disease
Richard A. Carleton	1931 - 2001	MD	anticipated	Brown Medical School	clinical studies of diet and smoking as cardiovascular disease risk factors
Susan M. Sieber	1942 - 2002	PhD	anticipated	National Cancer Institute	biochemical epidemiology and cancer
Henry Kamin	1920 - 1988	PhD	unknown	Duke University	biochemist
Jay P. Sanford	1928 - 1996	MD	anticipated	Univ. Texas Southwestern Medical School	internal medicine
George J. Schroepfer, Jr.	1932 - 1998	MD/PhD	sudden	Rice University	regulation of the formation and metabolism of cholesterol
James W. Prahl	1931 - 1979	MD/PhD	sudden	University of Utah	structural basis of the functions of human complement
Ira M. Goldstein	1942 - 1992	MD	anticipated	UCSF	pancreatitis, complement and lung injury
Milton H. Stetson	1943 - 2002	PhD	anticipated	University of Delaware	comparative endocrinology
David H.P. Streeten	1921 - 2000	MD/PhD	sudden	SUNY Upstate Medical University at Syracuse	thyroid and parathyroid hormones in hypertension
Victor D. Herbert	1927 - 2002	MD	anticipated	Harvard	biochemistry / clinical nutrition
Giovanni Di Chiro	1926 - 1997	MD	anticipated	NIH	interventional neuroradiology
John L. Doppman	1928 - 2000	MD	anticipated	NIH Clinical Center	radiologist
Arnold M. Seligman	1912 - 1976	MD	anticipated	Johns Hopkins University School of Medicine	chemistry
Robert H. Williams	1909 - 1979	MD	sudden	University of Washington	diabetes
Arthur T. Winfree	1942 - 2002	PhD	anticipated	University of Arizona	ecology and evolutionary biology
E. Carwile LeRoy	1933 - 2002	MD	sudden	University of South Carolina	rheumatology
Walsh McDermott	1901 - 1981	MD	sudden	Cornell University Medical College	public health and medicine
Lucien B. Guze	1928 - 1985	MD	sudden	UCLA	pathogenesis of experimental pyelonephritis
Zanvil A. Cohn	1926 - 1993	MD	sudden	Rockefeller University	macrophage in cell biology and resistance to infectious disease
Seymour Perry	1921 - 2000	MD	anticipated	Georgetown University Medical Center	evaluation of medical technology
Donald T. Witiak	1935 - 1998	PhD	sudden	University of Wisconsin	stereochemical studies of hypocholesterolemic agents
Kelly M. West	1925 - 1980	MD	sudden	University of Oklahoma	diabetes
Thomas A. McMahon	1943 - 1999	PhD	sudden	Harvard University	orthopedic biomechanics
Harold Weintraub	1945 - 1995	MD/PhD	anticipated	University of Washington/FHCRC	characterization and function of myod gene
Julio V. Santiago	1942 - 1997	MD	sudden	Washington University in St. Louis	role of social factors, lifestyle practices, and medication in the onset of type ii diabetes
John J. Wasmuth	1946 - 1995	PhD	sudden	University of California — Irvine	human-hamster somatic cell hybrids/localization of hnyington's disease gene
Anthony Dipple	1940 - 1999	PhD	sudden	NIH	metabolic activation and dna interactions of polycyclic aromatic hydrocarbon carcinogens
Fritz E. Dreifuss	1926 - 1997	MD	anticipated	University of Virginia School of Medicine	clinical investigations of childhood epilepsy
Robert D. Allen	1927 - 1986	PhD	anticipated	Dartmouth College	cytoplasmic rheology of motile cells
Demetrios Papahadjopoulos	1934 - 1998	PhD	sudden	UCSF	phospholipid-protein interactions, lipid vesicles, and membrane function
Julio H. Garcia	1933 - 1998	MD	sudden	Case Western Reserve University	neuroscience
Richard J. Wyatt	1939 - 2002	MD	anticipated	NIH	biochemistry of schizophrenia
Jurgen Steinke	1932 - 1973	MD	sudden	USC School of Medicine/Rancho Los Amigos Hospital	internal medicine/endoctinology
John P. Fox	1908 - 1987	MD/PhD	unknown	University of Washington	epidemiology
Kiichi Sagawa	1926 - 1989	MD/PhD	anticipated	Johns Hopkins University	modelling the mechanics of cardiac chamber contraction
Thoralf M. Sundt, Jr.	1930 - 1992	MD	anticipated	Mayo Clinic	surgical techniques for intracranial aneurysms
Ernest Bueding	1910 - 1986	MD	anticipated	Johns Hopkins School of Hygiene and Public Health	biochemistry and pharmacology
Lynn M. Wiley	1947 - 1999	PhD	sudden	University of California — Davis	morphogenesis in early mammalian embryos
Eleanor M. Saffran	1938 - 2002	PhD	anticipated	Temple University School of Medicine	cognitive deficits in brain-damaged patients
T. C. [Tao-Chiuh] Hsu	1917 - 2003	PhD	unknown	University of Texas MD Anderson Cancer Center	human and mammalian cytogenetics
Irving Kupfermann	1938 - 2002	PhD	anticipated	Columbia University	behavioral and neural analysis of learning in aplaysia
Vincent L. DeQuattro	1933 - 2001	MD	sudden	University of Southern California	cardiology
Abraham White	1908 - 1980	PhD	sudden	University of California	endocrinology
William F. Harrington	1920 - 1992	PhD	sudden	Johns Hopkins University School of Medicine	myosin thick filament structure and assembly
Mehdi Tavassoli	1933 - 1993	MD	anticipated	University of Mississippi Medical Center	hematopoietic stem cell purification and biology
William M. McKinney	1930 - 2003	MD	anticipated	Wake Forest University	neurology

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
Frank A. Beach	1911 - 1988	PhD	sudden	University of California, Berkeley	psychobiologist
Erwin Neter	1909 - 1983	MD	sudden	State University of New York at Buffalo	microbiology
Richard E. Bailey	1929 - 1972	MD	unknown	Case Western Reserve University	metabolism and diabetes
William B. Reed	1924 - 1976	MD	sudden	University of Southern California	dermatology
Ora M. Rosen	1935 - 1990	MD	anticipated	Sloan Kettering Institute for Cancer Research	cloning and characterization of gene for human insulin receptor
Donald J. Cohen	1940 - 2001	MD	anticipated	Yale University	tourette's syndrome and autism in children
Ernest G. Peralta	1959 - 1999	PhD	anticipated	Harvard University	signal transduction mechanisms of muscarinic receptors
Carl V. Moore	1908 - 1972	MD	sudden	Washington University in St. Louis	medicine
Richard M. Asofsky	1933 - 2000	MD	anticipated	National Institue of Allergy and Infectious Diseases	biomedical research
Harvey D. Preisler	1941 - 2002	MD	anticipated	Rush Medical College	clinical and biological studies of myeloid leukemias
James N. Davis	1939 - 2003	MD	sudden	SUNY HSC at Stony Brook	mechanisms underlying neuronal injury after brain ischemia
Henry C. Krutzsch	1942 - 2003	PhD	sudden	NIH	biochemistry
Solomon A. Berson	1918 - 1972	MD	sudden	Mount Sinai School of Medicine	diabetes and endocrinology
Earl W. Sutherland, Jr.	1915 - 1974	MD	unknown	Vanderbilt University	biochemist
Jack E. White	1921 - 1988	MD	anticipated	Howard University School of Medicine	epidemiology and treatment of cancer among african-americans
Christopher L. Longcope	1928 - 2003	MD	anticipated	UMASS	reproductive function and gonadal steroid dynamics
Eric Holtzman	1939 - 1994	PhD	sudden	Columbia University	cell biolgst
John G. Gambertoglio	1947 - 2001	PharmD	anticipated	UCSF	pharmacokinetics in healthy volunteers and subjects with renal insufficiency and on hemodialysis
Laird S. Cermak	1942 - 1999	PhD	anticipated	Boston University	psychological studies of memory and cognitive deficits related to chronic alcoholism
Richard J. Herrnstein	1930 - 1994	PhD	sudden	Harvard University	psychology
Thomas F. Gallagher	1905 - 1975	PhD	unknown	Albert Einstein College of Medicine	endocrinology
Koloman Laki	1909 - 1983	PhD	sudden	NFCF Regional Director	physical biochemistry
Grant W. Liddle	1921 - 1989	MD	sudden	Vanderbilt University	liddle's syndrome
Thomas K. Tatemichi	1952 - 1995	MD	anticipated	Columbia University College of Physicians and Surgeons	mechanisms and syndromes of dementia related to stroke
F. Blair Simmons	1930 - 1998	MD	sudden	Stanford University School of Medicine	development of a cochlear prothesis system for hearing loss
David S. Sigman	1939 - 2001	PhD	anticipated	UCLA	enzymology and gene targeting
Maurice S. Raben	1915 - 1977	MD	sudden	Tufts Medical School	endocrinology
Bruce S. Schoenberg	1942 - 1987	MD	anticipated	NIH	prevention and control of neurological disorders
Albert S. Kaplan	1917 - 1989	PhD	unknown	Vanderbilt University	microbiology
Peter A. Kollman	1944 - 2001	PhD	anticipated	UCSF	free energy perturbation calculations and their application to macromolecules
William T. Niemer	1911 - 1971	PhD	sudden	Creighton University	anatomy
Edward V. Evarts	1926 - 1985	MD	sudden	NIH	electrophysiological activity of in vivo neurons in waking and sleeping states
Ardie Lubin	1920 - 1976	PhD	anticipated	Naval Health Research Center	psychophysiology
Albert Segaloff	1917 - 1985	MD	sudden	Tulane University School of Medicine	hormonal treatment of advanced breast cancer
Janis V. Giorgi	1947 - 2000	PhD	anticipated	UCLA	cellular immunology of resistance to hiv
Benjamin E. Volcani	1915 - 1999	PhD	anticipated	University of California, La Jolla	microbiology
Sidney Futterman	1929 - 1979	PhD	anticipated	University of Washington	biochemistry of the retina and pigment epithelium
Leonard N. Horowitz	1947 - 1992	MD	anticipated	University of Pennsylvania School of Medicine	diagnosing and treatment of ventricular arrythmia
Alexander B. Gutman	1902 - 1973	MD/PhD	sudden	Mount Sinai School of Medicine	cancer research
Alvin Nason	1919 - 1978	PhD	unknown	Johns Hopkins University	biology
Susumu Hagiwara	1922 - 1989	PhD	sudden	UCLA	evolutionary and developmental properties of calcium channels in cell membranes
Sarah H. Broman	1927 - 1999	PhD	sudden	Office of Extramural Research, NINDS	research psychologist
Donald F. Summers	1934 - 2001	MD	anticipated	NIH	composition, assembly and replication of rna viruses
Patricia S. Goldman-Rakic	1937 - 2003	PhD	sudden	Yale University	development and plasticity of the primate frontal lobe
John W. Porter	1915 - 1984	PhD	unknown	University of Wisconsin	regulation of lipogenesis by insulin and glucagon
Leo K. Bustad	1920 - 1998	PhD	anticipated	Washington State University	radiation biology and physiology
Arend Bouhuys	1926 - 1979	MD/PhD	sudden	Yale University	epidemiology
David E. Green	1910 - 1983	PhD	anticipated	University of Wisconsin	molecular biology of membrane systems
Leo T. Samuels	1899 - 1978	PhD	unknown	University of Utah	biochemistry
Thomas G. Smith, Jr.	1931 - 1998	MD	sudden	NINDS	sensory physiology
William F. Caveness	1908 - 1981	MD	anticipated	NIH	authority on head injuries
Paul M. Aggeler	1911 - 1969	MD	anticipated	University of California	hemophilia
Michel M. Ter-Pogossian	1925 - 1996	PhD	sudden	Washington University School of Medicine	multislice pet scanning technology
Daniel S. Lehrman	1919 - 1972	PhD	sudden	Rutgers University	psychologist
W. Alden Spencer	1931 - 1977	MD	anticipated	Columbia University	physiology and neurology
Margaret O. Dayhoff	1925 - 1983	PhD	sudden	Georgetown University Medical Center	computer study of sequences of amino acids in proteins
D. Martin Carter	1936 - 1993	MD/PhD	sudden	Rockefeller University	susceptibility of pigment and cutaneous cells to dna injury by uv
Peter N. Magee	1921 - 2000	MD	unknown	Thomas Jefferson University	genetic basis of carconogenesis
Terry L. Thomas	1948 - 2002	PhD	anticipated	National Cancer Institute	radiation health effects
R. Gordon Gould	1910 - 1978	PhD	anticipated	Stanford University	internal medicine and cardiology
Peggy J. Copple	1934 - 1997	MD	sudden	University of Arizona College of Medicine	pediatrics and neurology
Ariel G. Loewy	1925 - 2001	PhD	sudden	Haverford College	biology
John N. Whitaker	1940 - 2001	MD	sudden	University of Alabama at Birmingham	neurology
Sidney Farber	1903 - 1973	MD	sudden	Harvard Medical School	cancer and biomedical research
Hans-Lukas Teuber	1916 - 1977	PhD	sudden	MIT	neuropsychologist
Sidney R. Cooperband	1931 - 1979	MD	unknown	Boston University Medical Center	internal medicine / hematology
Joram Heller	1934 - 1980	MD/PhD	anticipated	UCLA	biochemical and biophysical investigation of rhodopsin
Jack Schultz	1904 - 1971	PhD	sudden	University of Pennsylvania	genetics, tumor cells
Isadore Zipkin	1914 - 1973	PhD	anticipated	University of California- San Francisco	biochemistry / periodontology

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
Manfred M. Mayer	1916 - 1984	PhD	sudden	Johns Hopkins University School of Medicine	immunochemistry of the complement system
Timothy J. Regan	1924 - 2001	MD	anticipated	UMDNJ Newark	myocardial function and metabolism in chronic disease
Lester Baker	1930 - 2000	MD	anticipated	University of Pennsylvania School of Medicine/CHOP	clinical studies of type i diabetes control and complications
Alfred A. Smith	1928 - 1980	MD	unknown	New York Medical College	pharmacology
J. Kiffin Penry	1929 - 1996	MD	anticipated	Bowman Gray School of Medicine at Wake Forest University	epilepsy expert
Russell L. De Valois	1926 - 2003	PhD	sudden	University of California — Berkeley	brain mechanisms underlying color vision
Felix T. Rapaport	1929 - 2001	MD	sudden	SUNY Stony Brook	induction of unresponsiveness to allografts
Gerald P. Murphy	1934 - 2000	MD	sudden	Roswell Park Cancer Institute/SUNY Buffalo	detection, immunotherapy, and prognostic indicators of prostate cancer
D. Michael Gill	1940 - 1990	PhD	sudden	Tufts University	biochemistry of cholera toxin and other pathogenic toxins
Robert C. Schlant	1929 - 2002	MD	anticipated	Emory University	cardiology
Klaus Schwarz	1914 - 1978	MD	sudden	UCLA School of Medicine	biological chemistry
Per F. Scholander	1905 - 1980	MD/PhD	sudden	Scripps Institute of Oceanogrpahy in La Jolla, California	physiology
Donald J. Magilligan, Jr.	1929 - 1989	MD	sudden	Henry Ford Health Sciences Center	natural history and limitations of porcine heart valves
Leslie A. Stauber	1907 - 1973	PhD	sudden	Rutgers University	zoology
Dominick E. Gentile	1932 - 1997	MD	sudden	St Joseph Hospital, Orange, California	nephrology
Sukdeb Mukherjee	1946 - 1995	MD	sudden	Medical College of Georgia, Augusta	psychiatry
George K. Smelser	1908 - 1973	PhD	sudden	Columbia University	ophtalmology
Sheldon D. Murphy	1933 - 1990	PhD	anticipated	University of Washington	toxicology
James S. Seidel	1943 - 2003	MD/PhD	sudden	Harbor-UCLA Medical Center	clinical studies in pediatric life support and cardiopulmonary resuscitation
John Gibbon	1934 - 2001	PhD	anticipated	Columbia University	cns functions underlying the interval time sense in animals and humans
Edwin H. Beachey	1934 - 1989	MD	anticipated	University of Tennessee at Memphis	chemistry and immunology of streptococcal m proteins
John M. Eisenberg	1946 - 2002	MD	anticipated	Georgetown University	internal medicine
Merton F. Utter	1917 - 1980	PhD	sudden	Case Western Reserve University School of Medicine	structure and function of pep carboxykinase isozymes
Cornelia P. Channing	1938 - 1985	PhD	anticipated	University of Maryland School of Medicine	physiology
Jane Pitt	1938 - 2003	MD	anticipated	Columbia University College of Physicians and Surgeons	perinatal transmission of hiv and retroviral infections
Norbert Freinkel	1926 - 1989	MD	sudden	Northwestern University	metabolic regulation in normal and diabetic pregnancies
Richard L. Lyman	1927 - 1975	PhD	anticipated	UC Berkeley	biochemistry
Dexter French	1918 - 1981	PhD	unknown	Iowa State University	biochemistry and biophysics
M. Powell Lawton	1923 - 2001	PhD	anticipated	Philadelphia Geriatric Center	studies of mental health, quality of life, and caregiving of the elderly
Mortimer M. Elkind	1922 - 2000	PhD	anticipated	Colorado State University	cell radiation response of cultured mammalian cells
Samuel Schwartz	1916 - 1997	MD	anticipated	University of Minnesota	cancer research
Cornelius A. Tobias	1918 - 2000	PhD	anticipated	University of California — Berkeley	biological effects of cosmic rays and other ionizing radiation
Samuel B. Guze	1923 - 2000	MD	anticipated	Washington University School of Medicine	neurobiology, genetics, and epidemiology of alcoholism
Sidney P. Colowick	1916 - 1985	PhD	unknown	Vanderbilt University	enzymatic oxidation and phosphorylation
Harold A. Baltaxe	1931 - 1985	MD	sudden	University of California — Davis	development of new coronary angiographic techniques
Lee A. Lillard	1943 - 2000	PhD	sudden	University of Minnesota Retirement Research Center	demography
Bruce M. Achauer	1943 - 2002	MD	sudden	University of California — Irvine	non-invasive methods to assess the depth of burn wounds
S. Smith Stevens	1906 - 1973	PhD	sudden	Harvard University	psychophysics
S. Bernard Wortis	1904 - 1969	MD	sudden	New York University School of Medicine	psychiatrist and neurologist
Charles K. Friedberg	1905 - 1972	MD	sudden	Mount Sinai Hospital	pathology, biochemistry
Benjamin Alexander	1908 - 1978	MD	unknown	NY Blood Center	medicine
Lawrence H. Piette	1932 - 1992	PhD	anticipated	Utah State University	electron spin resonance spectroscopy
Robert A. Mendelson, Jr.	1941 - 2001	PhD	anticipated	UCSF	molecular mechanism of muscle contraction
Roger M. Brown	1941 - 2002	PhD	sudden	NIDA (Neuro-Immune Disease Alliance)	neuroscience
Edward C. Heath	1930 - 1985	PhD	unknown	University of Iowa	biochemist
Elizabeth M. Smith	1939 - 1997	PhD	anticipated	Washington University School of Medicine	psychiatric problems among disaster survivors
David F. Waugh	1915 - 1984	PhD	unknown	MIT	biochemistry
Allan Beigel	1940 - 1996	PhD	anticipated	University of Arizona	psychiatry and mental healh policy
Leo J. Neuringer	1928 - 1993	PhD	anticipated	MIT	nmr studies of normal and transformed cell membranes
Arnost Kleinzeller	1914 - 1997	MD/PhD	anticipated	University of Pennsylvania Medical School	physiology
Norman R. Davidson	1916 - 2002	PhD	sudden	California Institute of Technology	physical chemistry of nucleic acids
Lawrence D. Jacobs	1938 - 2001	MD	anticipated	SUNY Buffalo	recombinant b interferon as treatment for multiple sclerosis
Abraham Worcel	1938 - 1989	MD	sudden	University of Rochester	cell biologst
Albert H. Coons	1912 - 1978	MD	sudden	Harvard Medical School	bacteriology and immunology
Monroe E. Wall	1916 - 2002	PhD	sudden	Research Triangle Institute	isolation and chemistry of plant antitumor agents
William B. Kinter	1926 - 1978	PhD	unknown	Mount Desert Island Biological Lab	physiology
Sidney Riegelman	1921 - 1981	PhD	sudden	UCSF	intersubject variation in first pass effect of drugs
Charles L. Wittenberger	1930 - 1987	PhD	sudden	Intramural Research Program of the NIDR	microbiology
Wigbert C. Wiederholt	1931 - 2000	MD	anticipated	UCSD	age related neurodegenerative diseases in micronesia
Donnell D. Etzwiler	1927 - 2003	MD	anticipated	University of MInnesota	pediatrician
Emanuel M. Bogdanove	1925 - 1979	PhD	sudden	Medical College of Virginia at Richmond	physiology
Griff T. Ross	1921 - 1985	PhD	anticipated	NIH	endocrinologist
James C. Steigerwald	1935 - 1988	MD	sudden	University of Colorado School of Medicine	internal medicine / rheumatology
J. Herbert Conway	1904 - 1969	MD	sudden	Cornell University Medical College	surgery
Barbara J. Lowery	1938 - 2002	PhD	anticipated	University of Pennsylvania	understanding stress responses of people who were physically ill
Larry C. Clark	1948 - 2000	PhD	anticipated	University of Arizona	nutritional prevention of cancer
Irwin M. Weinstein	1926 - 2002	MD	sudden	Cedars-Sinai Medical Center	hematology
Charles E. Huggins	1930 - 1990	MD	anticipated	Harvard University Medical School	human blood storage
Abraham M. Lilienfeld	1920 - 1984	MD	sudden	Johns Hopkins University School of Public Health	epidemiological methods for the study of chronic diseases

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
Daniel A. Brody	1915 - 1975	MD	sudden	University of Tennessee	biophysics
Frank Restle	1927 - 1980	PhD	sudden	Indiana University at Bloomington	psychology
Aubrey Gorbman	1914 - 2003	PhD	anticipated	University of Washington in Seattle	endocrinology
Henry A. Schroeder	1906 - 1975	MD	unknown	Dartmouth Medical School	physiology
Josiah Brown	1923 - 1985	MD	sudden	UCLA	biochemical studies of lipid and carbohydrate metabolism
Alexander S. Wiener	1907 - 1976	MD	anticipated	New York University	genetics and biometrics
David H. Blankenhorn	1924 - 1993	MD	anticipated	USC Keck School of Medicine	control of risk factors in atherosclerosis
Joseph Cochin	1916 - 1985	MD/PhD	anticipated	Boston University Medical Center	pharmacology
Thomas F. Necheles	1933 - 1984	MD/PhD	sudden	NCI	biology and infectious disease
Cyrus Levinthal	1922 - 1990	PhD	anticipated	Columbia University College of Physicians and Surgeons	colinearity of genes and proteins, and the nature of messenger rna
Marshall R. Urist	1914 - 2001	MD	anticipated	UCLA	inductive substrates of tooth and bone formation
Victor A. Gilbertsen	1924 - 1990	MD	unknown	University of Minneosta	surgical oncology
Paul A. Bunn	1914 - 1970	MD	sudden	University of Michigan	tuberculosis
John L. Kemink	1949 - 1992	MD	sudden	University of Michigan, Ann Arbor	otolaryngology
Tai-Shun Lin	1939 - 1994	PhD	anticipated	Yale University School of Medicine	synthesis and development of nucleoside analogs as antiviral and anticancer compounds
Norman Geschwind	1926 - 1984	MD	sudden	Harvard Medical School/Beth Israel Medical Center	relationship between the anatomy of the brain and behavior
Lucien J. Rubinstein	1924 - 1990	MD	sudden	University of Virginia School of Medicine	differentiation and stroma-induction in neural tumors
Marion I. Barnhart	1921 - 1985	PhD	sudden	Wayne State University School of Medicine	blood disorders
Roy D. Schmickel	1936 - 1990	MD	sudden	University of Pennsylvania	isolation and characterization of human ribosomal dna
Mones Berman	1920 - 1982	PhD	anticipated	National Cancer Institute	mathematical biology
Edgar C. Henshaw	1929 - 1992	MD	sudden	University of Rochester	intermediary metabolism in animals and in man
Philip R. Kimbel	1925 - 1990	MD	unknown	University of Pennsylvania School of Medicine	causes of emphysema and other pulmonary diseases
Edward W. Dempsey	1911 - 1975	PhD	sudden	Columbia University	anatomy-cytology (microbiology)
Kehl Markley, 3rd	1923 - 1979	MD	sudden	NIH / NIAMDD	burn treatment specialist
Arnold Lazarow	1916 - 1975	MD/PhD	sudden	University of Minnesota	anatomy
Shu-Ren Lin	1936 - 1979	MD	sudden	University of Rochester	imaging studies of cerebral blood flow after cardiac arrest
Jessica H. Lewis	1917 - 2003	MD	sudden	University of Pittsburgh	physician
Alberto DiMascio	1928 - 1978	PhD	unknown	Tufts University	psychiatry and pharmacology
G. Harrison Echols, Jr.	1933 - 1993	PhD	anticipated	University of California — Berkeley	genetic and chemical studies of phage lambda development
V. Everett Kinsey	1909 - 1978	PhD	sudden	Institute of Biological Sciences at Oakland University	ophthalmology (eye)
Leah M. Lowenstein	1931 - 1984	MD/PhD	unknown	Jefferson Medical College	internal medicine / nephrology
Eugene C. Jorgensen	1923 - 1981	PhD	sudden	UCSF	structure/activity relationships of compounds related to thyroxin
Ronald E. Talcott	1947 - 1984	PhD	sudden	UCSF	carboxylesterases of toxicologic significance
William J. Bowen	1911 - 1970	PhD	sudden	National Institute of Arthritis and Metabolic Diseases	biophysics
Wendell M. Stanley	1904 - 1971	PhD	sudden	University of California	biochemist
Herman K. Hellerstein	1916 - 1993	MD	anticipated	Case Western Reserve University	cardiology
Thaddeus S. Danowski	1914 - 1987	MD	sudden	University of Pittsburgh	research medicine
Thomas S. Whitecloud, 3rd	1940 - 2003	MD	sudden	Tulane University	orthopaedics
Cesare G. Tedeschi	1904 - 1974	MD	unknown	Framingham Union Hospital	pathology
John A. Gronvall	1931 - 1990	MD	sudden	Veterans Administration in Washington,D.C.	pathology
Michale E. Keeling	1942 - 2003	MD	sudden	University of Texas	veterinary science
Philip G. Weiler	1941 - 1991	MD	anticipated	University of California — Davis	coronary heart disease & stroke in the elderly
Harold P. Morris	1900 - 1982	PhD	sudden	Howard University College of Medicine	cancer research
Howard J. Eisen	1942 - 1987	MD	sudden	National Institute of Child Health and Human Development	physician
Paul Margolin	1923 - 1989	PhD	sudden	City College of New York	genetics researcher
William W. Montgomery	1923 - 2003	MD	anticipated	Harvard Medical School	medicine and otolaryngology
Sandra A. Daugherty	1934 - 2000	MD/PhD	anticipated	University of Nevada, Reno	epidemiologist
Guillermo H. Pacheco	1931 - 1974	PhD	anticipated	National Institute of Allergy and Infectious Diseases / NIH	microbiology
Michael Doudoroff	1911 - 1975	PhD	anticipated	University of California, Berkeley	bacteriology and immunology
Frederick S. Philips	1916 - 1984	PhD	anticipated	Sloan Kettering Institute for Cancer Research	pharmacological properties of chemotherapeutic agents and chemical carcinogenesis
Ernst Simonson	1898 - 1974	MD	sudden	University of Minnesota	cardiology and physiology
Jacob W. Dubnoff	1909 - 1972	PhD	anticipated	University of Southern California	neurology
Dorothy T. Krieger	1927 - 1985	MD	anticipated	Mount Sinai School of Medicine	cns-pituitary-adrenal interactions
David W. Fulker	1937 - 1998	PhD	anticipated	University of Colorado at Boulder	adoption studies of development in middle childhood
Charles W. Mays	1930 - 1990	PhD	anticipated	National Cancer Institute	radiobiology
Charles H. Rammelkamp, Jr.	1911 - 1981	MD	sudden	Case Western Reserve University School of Medicine	early studies on the clinical application & mechanism of action of antimicrobials
Gustavo Cudkowicz	1927 - 1982	MD	sudden	SUNY Buffalo	controls of proliferation specific for leukemias
Herschel L. Roman	1914 - 1989	PhD	sudden	University of Washington	geneticist
Laurence M. Sandler	1929 - 1987	PhD	sudden	University of Washington School of Medicine	cytogenetics of meiosis and development in drosophila
William D. Nunn	1943 - 1986	PhD	sudden	University of California — Irvine	regulation of fatty acid/acetate metabolism in e. coli
William Likoff	1912 - 1987	MD	unknown	Hahnemann Medical College	cardiology
Robert A. Cooper, Jr.	1932 - 1992	MD	sudden	University of Rochester Cancer Center	pathologist
Caroline T. Holloway	1937 - 1998	PhD	sudden	NIH / Center for Accelerator Mass Spectrometry at Lawrence Livermore National Laboratory	biomedical research
Marilyn Bergner	1933 - 1992	PhD	anticipated	Johns Hopkins School of Hygiene and Public Health	health policy and management
Wylie J. Dodds	1934 - 1992	MD	anticipated	Medical College of Wisconsin	esophageal motor function in health and disease
Mearl F. Stanton	1922 - 1980	MD	anticipated	National Cancer Institute	pathology
Edwin D. Murphy	1917 - 1984	MD	unknown	former head of the Research Unit on Gynecologic Pathology, NCI	experimental pathology
Arthur E. Martell	1916 - 2003	PhD	unknown	Texas A & M	chemistry
Peter W. Neurath	1923 - 1977	PhD	sudden	Tufts University	physics in medicine

Investigator Name	Lifespan	Degree	Death Type	Institution at Time of Death	Scientific Domain
Alfred S. Evans	1917 - 1996	MD	anticipated	Yale Medical School	physician
Nathaniel A. Young	1939 - 1979	MD	sudden	National Cancer Institute	oncology and molecular pathology
Hermann Rahn	1912 - 1990	PhD	anticipated	University of Buffalo School of Medicine	physiologist
George N. Wise	1915 - 1974	MD	sudden	Albert Einstein College of Medicine	ophthalmology
John C. Cassel	1921 - 1976	MD	anticipated	University of North Carolina School of Public Health	epidemiology
Alan S. Morrison	1943 - 1992	PhD	anticipated	Brown Medical School	hormones in the epidemiology of prostatic hyperplasia
David Zeaman	1921 - 1984	PhD	unknown	University of Connecticut	psychology
Jerry D. Niswander	1930 - 1984	MD	anticipated	National Institute of Dental Research	genetics of oral and facial disorders
Lauran D. Harris	1927 - 1987	MD	anticipated	Boston University Medical Center	medicine
Kiertisin Dharmsathaphorn	1950 - 1990	MD	anticipated	UCSD	intestinal secretory mechanisms and antidiarrheal drugs
Harvey C. Knowles, Jr.	1915 - 1984	MD	anticipated	University of Cincinnati/Children's Hospital	clinical studies of gestational diabetes
Ernest Cotlove	1920 - 1970	MD	sudden	NIH	clinical pathology
Chandler McC. Brooks	1905 - 1989	PhD	sudden	State University of New York	neurophysiology and cardiology
James D. Hardy	1904 - 1985	PhD	anticipated	University of Mississippi Medical Center	transplant surgery
George V. Taplin	1910 - 1979	MD	anticipated	UCLA	radioactive albumin macroaggregates for the detection of pulmonary embolism
Maurice Landy	1913 - 1993	PhD	unknown	NIH	genetic control of immune responsiveness
Ronald D. Fairshter	1942 - 1988	MD	anticipated	University of California — Irvine	clinical studies in chronic obstructive pulmonary disease
Elizabeth A. Bates	1947 - 2003	PhD	anticipated	UCSD	cross-linguistic studies of language development, processing and breakdown in aphasia
Peter Kellaway	1920 - 2003	PhD	anticipated	Baylor College of Medicine	clinical investigations of childhood epilepsy
Richard E. Weitzman	1943 - 1980	MD	anticipated	Harbor-UCLA Medical Center	arginine vasopressin metabolism
Grace A. Goldsmith	1904 - 1975	MD	unknown	Tulane School of Public Health and Tropical Medicine	nutrition and dietary diseases
Walter E. Brown	1918 - 1993	PhD	unknown	American Dental Association Health Foundation	chemistry of calcium phosphates
Max Halperin	1917 - 1988	PhD	anticipated	George Washington University	biostatistics
Joseph Stokes, 3rd	1924 - 1989	MD	anticipated	Boston University School of Medicine	epidemiological studies of coronary heart disease
William J. Meyers	1933 - 1970	PhD	anticipated	University of Louisville	psychology and physiology
Philip R.A. May	1920 - 1986	MD	anticipated	UCLA	controlled clinical studies of schizophrenia
JoAnn E. Franck	1950 - 1992	PhD	anticipated	University of Washington School of Medicine	hippocampal damage as a cause of epilepsy
Stephen W. Kuffler	1913 - 1980	MD	sudden	Harvard University	neurobiology
Pokar M. Kabra	1942 - 1990	PhD	sudden	UCSF	application of liquid chromatography to therapeutic drug monitoring
David D. Rutstein	1909 - 1986	MD	sudden	Harvard University Medical School	physician
Nathan W. Shock	1906 - 1989	PhD	anticipated	NIH	gerontology
Howard E. Freeman	1929 - 1992	PhD	sudden	UCLA	sociology
Judith G. Pool	1919 - 1975	PhD	anticipated	Stanford University	physiologist
Lois W. Tice	1934 - 1985	MD	sudden	NIADDK (National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases)	cell biology
Robert Thompson	1927 - 1989	PhD	anticipated	University of California — Irvine	neural systems subserving learning and memory
Robert B. Woodward	1917 - 1979	PhD	sudden	Harvard University	organic chemistry
Morton J. Hamburger	1907 - 1970	MD	sudden	University of Cincinnati College of Medicine	infectuous deseases
Stanley R. Kay	1946 - 1990	PhD	sudden	Albert Einstein College of Medicine	symptoms and diagnostic tests of schizophrenia
John P. Glynn	1932 - 1971	PhD	sudden	National Cancer Institute	immunology, tumor biology
Alex B. Novikoff	1913 - 1987	PhD	unknown	Albert Einstein College of Medicine of Yeshiva University	histochemical studies of the golgi apparatus
Ann L. Brown	1943 - 1999	PhD	sudden	University of California, Berkeley Graduate School of Education	educational theorist
Edward A. Steinhaus	1914 - 1969	PhD	sudden	University of California, Irvine	biology
Morris B. Bender	1905 - 1983	MD	sudden	Mount Sinai Hospital	neurology
Janine André-Schwartz	1931 - 1995	MD	anticipated	Tufts University	immunology
Alexander D. Langmuir	1910 - 1993	MD	anticipated	Johns Hopkins University School of Hygiene and Public Health	epidemiology
Harry A. Feldman	1914 - 1985	MD	anticipated	State University of New York Upstate Medical Center College of Medicine	physician, epidemiologist
Eli Chernin	1924 - 1990	PhD	sudden	Harvard School of Public Health	parasitologist
George E. Murphy	1918 - 1987	MD	anticipated	Cornell University Medical College	pathologist
Arnold M. Mordkoff	1936 - 1971	PhD	sudden	New York University	psychology
Bernard G. Greenberg	1919 - 1985	PhD	anticipated	University of North Carolina, School of Pubic Health	biostatistics
Edward W. Purnell	1928 - 1993	MD	anticipated	Case Western Reserve University	ophthalmology
Eugenia Spanopoulou	1960 - 1998	PhD	sudden	Mount Sinai School of Medicine	cancer research
Jack Metcoff	1917 - 1994	MD	unknown	Chicago Medical School	nephrologist, nutritionist
John E. Howard	1902 - 1985	MD	sudden	Johns Hopkins Hospital	endocrinology
Mindel C. Sheps	1913 - 1973	MD	anticipated	University of North Carolina at Chapel Hill	biostatistics and demography
James L. Lehr	1940 - 1989	MD	unknown	University of Chicago	physician
A. Louis McGarry	1929 - 1985	MD	anticipated	Nassau County Department of Mental Health	psychiatry
Lester R. Dragstedt	1893 - 1975	MD/PhD	sudden	University of Chicago	surgery
Roberta D. Shahin	1953 - 1997	PhD	sudden	Center for Biologics Evaluation and Research	immunology
Donovan J. Thompson	1919 - 1991	PhD	sudden	University of Washington	medicine and biostatistics
William G. Dauben	1919 - 1997	PhD	sudden	University of California — Berkeley	ultraviolet irradiation of natural products

Online Appendix References

Azoulay, Pierre, Joshua S. Graff Zivin, and Gustavo Manso. 2011. "Incentives and Creativity: Evidence from the Academic Life Sciences." *The RAND Journal of Economics* **42**(3): 527-554.

Azoulay, Pierre, Andrew Stellman, and Joshua Graff Zivin. 2006. "PublicationHarvester: An Open-Source Software Tool for Science Policy Research." R*esearch Policy* **35**(7): 970-974.

Blackwell, Matthew, Stefano M. Iacus, Gary King, and Giuseppe Porro. 2009. "CEM: Coarsened Exact Matching in Stata." *The Stata Journal* **9**(4): 524-546.