

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Medical Care Output and Productivity

Volume Author/Editor: David M. Cutler and Ernst R. Berndt, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-13226-9

Volume URL: <http://www.nber.org/books/cutl01-1>

Publication Date: January 2001

Chapter Title: The Allocation of Publicly Funded Biomedical Research

Chapter Author: Frank R. Lichtenberg

Chapter URL: <http://www.nber.org/chapters/c7642>

Chapter pages in book: (p. 565 - 590)

The Allocation of Publicly Funded Biomedical Research

Frank R. Lichtenberg

In the last century, the average health of the American people has improved dramatically. The mean life expectancy of Americans has increased almost twenty years, or two years per decade,¹ since the turn of the century. Just from 1979 to 1988, the age-adjusted mortality rate declined 7.2 percent.

An important part of this enormous progress in health (which is scarcely reflected in our national accounts) is probably due to large private and public investments in biomedical research. In 1993, health R&D accounted for 18 percent of total U.S. R&D expenditure. Health R&D expenditures, by source of funding, are shown in table 15.1.

The National Institutes of Health (NIH) administer about 80 percent of federal health R&D. The NIH is made up of twenty-one institutes and centers, each with a mission and a separate, annual budget established by Congress. The institutes and centers are listed in table 15.2, along with the year in which each was established and its fiscal year 1998 budget obligations.² NIH does not want people to be misled by the names of the institutes; it points out that “research on any disease is not confined to

Frank R. Lichtenberg is the Courtney C. Brown Professor of Business at Columbia University and a research associate of the National Bureau of Economic Research.

The author is grateful to David Cutler and Ernie Berndt for helpful comments on a previous draft, and to the American Enterprise Institute and the Center for Economic Studies at the University of Munich for financial support. He is responsible for any errors.

1. “Buy ten, get two free,” could be a fair, if crass, marketing slogan for U.S. health progress.

2. Older institutes and centers tend to receive significantly more funding than younger ones: There is a strong positive correlation ($r = .636$) between an institute or centers age in 1998 and (the log of) its budget obligations.

Table 15.1 Health R&D Expenditures

Source of Funding	1993 Health R&D Funds
Federal	12,051
State and local	2,054
Industry	15,711
Private nonprofit	1,215
All sources	31,032

Source: National Center for Health Statistics 1995, table 132.

Note: Figures are in millions of dollars.

Table 15.2 NIH Institute and Centers

	Year Established	1998 Obligations (million \$)
National Cancer Institute	1937	2547
National Institute of Mental Health	1946	750
National Heart, Lung, and Blood Institute	1948	1,531
National Institute of Allergy and Infectious Diseases	1948	1,352
National Institute of Dental Research	1948	209
National Institute of Diabetes and Digestive and Kidney Diseases	1950	874
National Institute of Neurological Disorders and Stroke	1950	781
National Center for Research Resources	1956	454
National Institute of Child Health and Human Development	1963	675
National Institute of General Medical Sciences	1963	1,066
National Institute of Environmental Health Sciences	1966	330
National Eye Institute	1968	356
National Library of Medicine	1968	161
John E. Fogarty International Center	1968	28
National Institute on Alcohol Abuse and Alcoholism	1970	227
National Institute on Aging	1974	519
National Institute on Drug Abuse	1974	527
National Institute of Arthritis and Musculoskeletal and Skin Diseases	1986	275
National Institute of Nursing Research	1986	64
National Institute on Deafness and Other Communication Disorders	1988	201
National Human Genome Research Institute	1989	218
Total		13,145

Source: <http://www.nih.gov/welcome/almanac/index.html>.

one Institute, and no Institute is dedicated to a single disease. An Institute's budget is an inadequate measure of support for research on specific diseases. Research into many diseases is often carried on in several Institutes simultaneously, e.g., several Institutes are supporting research on Alzheimer's disease."

While the NIH focuses much of its research on combating specific diseases, and much of its funding supports research projects that are of obvious relevance to specific diseases, the NIH also places a high priority on funding basic research. These basic research projects may appear initially to be unrelated to any specific disease, but might prove to be a critical turning point in a long chain of discoveries leading to improved health. Each of the NIH institutes supports basic research likely to advance particular areas of science that might prove relevant to clinical problems important to that institute's mission. By supporting disease-related and basic research projects simultaneously, the NIH seeks to achieve both near-term improvements in the diagnosis, treatment, and prevention of specific diseases and long-term discoveries in basic science that in time will produce great advances in our ability to understand, treat, and prevent disease or delay its onset.

In this paper I develop a simple theoretical model of the allocation of the applied component of public biomedical research expenditure—the approximately 50 percent of expenditure that is of direct, near-term relevance to specific diseases—and present some empirical evidence about the determinants of this allocation. The implications of the theoretical model are consistent with government officials' descriptions of the allocation process: The structure of expenditure should depend upon research productivity (or "scientific opportunity") as well as on public health need, that is, the societal and economic burden of the disease/condition. Although we lack, at this point, useful indicators of research productivity (i.e., of the cost of achieving research advances), we have a number of measures of disease burden (i.e., of the benefit of achieving these advances).³ Analysts of technical change typically have data on neither the costs nor the benefits of technical advance. Failure to measure research productivity will not necessarily bias my estimates; if it does, it seems likely to bias them toward zero.

The paper is organized as follows. In the next section I develop the simple model of public research expenditure allocation. I rely on three types and sources of data to estimate the parameters of the model: data on research activity derived from NIH's CRISP (Computerized Retrieval of Information on Scientific Projects) database, premature mortality data

3. The disease burden is the *potential* benefit (ignoring all comorbidities), not the actual benefit.

from the Vital Statistics–Mortality Detail file, and data on chronic condition prevalence and severity from the National Health Interview Survey. These are discussed in section 15.2. Preliminary estimates are presented in section 15.3, and a summary is provided in section 15.4.

15.1 A Simple Model of the Determinants of Research Expenditure at the Disease Level

To motivate the discussion and develop a few intuitions, I write down the simplest possible model of research funding allocation. This model is based on the following extremely strong assumptions (some of which are relaxed below): (1) there are only two diseases; (2) the number of people suffering from the two diseases, N_1 and N_2 , is exogenous; (3) the average severity of the two diseases is identical; (4) the probability P_i of finding a cure for disease i ($i = 1, 2$) is a concave (deterministic) function of research funding for that disease, X_i ; $P_i = X_i^\alpha$, where $0 < \alpha < 1$;⁴ (5) the effect of funding on the probability of finding a cure is the same across diseases; and (6) the total research budget $X = X_1 + X_2$ is fixed.

Suppose that policymakers attempt to maximize the (expected) total number of people cured of both diseases subject to the budget constraint,⁵ that is, they choose X_1 to maximize

$$\begin{aligned} (1) \quad J^* &= N_1 P_1 + N_2 P_2 \\ &= N_1 X_1^\alpha + N_2 X_2^\alpha \\ &= N_1 X_1^\alpha + N_2 (X - X_1)^\alpha. \end{aligned}$$

The first-order condition implies that relative funding of research on the two diseases should satisfy

$$(2) \quad \ln(X_1/X_2) = [1/(1 - \alpha)] \ln(N_1/N_2).$$

Research funding should increase with disease incidence: for example, $X_1 > X_2$ if $N_1 > N_2$. This is because the benefit of discovering a cure for the disease is proportional to its incidence, but the cost is independent of incidence. Moreover the elasticity of funding with respect to incidence should exceed unity: if disease 1 is twice as prevalent as disease 2, research funding for disease 1 should be more than twice as great as research fund-

4. Viscusi (1995, 3) notes that “in the case of biomedical research, the typical outcome will be a change in societal risk levels induced by the biomedical research outcomes.”

5. I assume for simplicity that federal policymakers do not pay attention to biomedical R&D funded by other sources; in other words, they are not merely trying to “fill gaps” in other research, nor do they consider the potential impact of public R&D on other research activity. Toole (2000), however, presents evidence that suggests that public biomedical research may have a significant, albeit very delayed, impact on private drug discovery.

ing for disease 2.⁶ Equalizing research expenditure per victim across diseases would be inefficient.

One could generalize this model to the case of $I > 2$ diseases, to obtain $I - 1$ equilibrium conditions of the form

$$(3) \quad \ln X_i = \text{constant} + [1/(1 - \alpha)] \ln N_i,$$

($i = 1, 2, \dots, I - 1$). Given cross-sectional or panel data on research funding and incidence by disease, one could estimate equation (3) to test the hypothesis of diminishing returns to research funding at the disease level and to estimate the parameter α . But this simple model can and should be extended in at least two directions: We should allow for multiple indicators of incidence and for differences in research productivity (scientific opportunity) across diseases.

15.1.1 Multiple Indicators of Incidence

As the director of NIH says, a given disease imposes a number of different kinds of burden on society, and “policy makers will need to consider the relative importance or weight to be placed on each criteri[on] when assessing the overall societal burden imposed by each disease.” While the NIH has indicated interest in determining how to measure public health burden, it has also expressed uncertainty about how to do so. I now outline a procedure for doing this.⁷ Then I will perform empirical analysis to ascertain how close the actual allocation of research resources is to the allocation that is optimal, according to my framework. The answer appears to be “pretty close.”

Suppose that the overall burden of a disease is perceived by policymakers to be a function of K attributes of the disease: $N_i \equiv f(A1_i, A2_i, \dots, AK_i)$ where, for example, $A1$ is the number of deaths, $A2$ is the number of bed-disability days, $A3$ is the number of hospital stays, and so forth. Further suppose that the functional form of this relationship is

$$(4) \quad \ln N_i = \beta_1 \ln A1_i + \beta_2 \ln A2_i + \dots + \beta_K \ln AK_i,$$

where $\sum_k \beta_k = 1$. The term β_k reveals the relative “weight” assigned by policymakers to attribute k in the determination of overall disease burden. Substituting equation (4) into equation (3),

6. In reality, finding a cure for one of the diseases may increase the probability of suffering at a future date from the other disease. Development of a richer model to account for this and other complications is a challenging task. Given the stark simplicity of my model, it may be best to view it as a set of organizing principles that can be used to interpret the allocative process, rather than as a theory.

7. Recently a National Academy of Sciences panel looked at priority setting at NIH, and recommended using a number of measures to measure burden of illness in a fashion similar to what I propose. See Institute of Medicine (1998).

$$(5) \quad \ln X_i = \text{constant} + [1/(1 - \alpha)] (\beta_1 \ln A1_i + \beta_2 \ln A2_i + \dots + \beta_k \ln AK_i),$$

Estimation of equation (5) would provide estimates of these (“revealed preference”) weights as well as of the technological parameter α . They would indicate the relative weight given to mortality and bed-disability days, for example.

Since disease outcome and incidence data are available by demographic group, we can also make inferences about weights associated with different demographic groups.⁸ For example, let us define “adjusted” bed-disability days $A2^* = A2YOUNG + (1 + \theta) A2OLD$, where $A2YOUNG$ and $A2OLD$ denote bed-disability days of young and old people, respectively. If policymakers’ evaluation of the marginal burden of the two groups’ bed-disability days differs, θ will differ from zero. This parameter can be estimated by replacing $A2$ by $A2^*$ in equation (5).

15.1.2 Differences in Research Productivity (Scientific Opportunity) across Diseases

The preceding model is based on the assumption that the effect of funding on the probability of finding a cure is the same across diseases. This assumption is clearly unrealistic, and it is desirable to relax it.⁹ We can modify the cure-probability equation to include a disease-specific research productivity parameter π_i ; $P_i = \pi_i X_i^\alpha$. The objective function policymakers seek to maximize is now $J^* = N_1 P_1 + N_2 P_2 = N_1 \pi_1 X_1^\alpha + N_2 \pi_2 X_2^\alpha$, and the optimal expenditure on research on disease i is now

$$(6) \quad \ln X_i = \text{constant} + [1/(1 - \alpha)] \ln N_i + [1/(1 - \alpha)] \ln \pi_i.$$

The research-productivity parameters i enter the objective function and the optimal expenditure equation in the same way as the disease incidence measures N_i . Research expenditure should be an increasing function of scientific opportunity as well as of disease burden. This implication is consistent with the views expressed by government officials: “It is vital that the allocation of medical research dollars takes into account several factors, including scientific opportunity, public health need, gaps in knowledge, as well as societal and economic burden of the disease/condition.”¹⁰

8. NIH officials acknowledge that “research funding decisions will also reflect concerns about equity among groups of potential beneficiaries of the research as defined in terms of age, sex, and ethnic origin. Certain criteria favor one group over another. For example, mortality rates and measures of the impact on functioning may favor the elderly whereas measures of economic impact, such as lost productivity, would favor younger citizens” (NIH Director Varmus’s responses to questions from Senator Slade Gordon, Labor, HHS, Education Subcommittee Hearing, NIH appropriations for FY 1996, 18 May 1995).

9. Henderson and Cockburn (1996) have studied the determinants of research productivity of pharmaceutical firms, using patents and scientific papers as measures of research output.

10. Office of Science Policy, NIH Response to Congressional Questions, June 1996. Garber and Romer (1995) also argue that “federal policy toward research and development should

I believe that the CRISP data can eventually be exploited to obtain indicators of (changes in) the relative productivity of research on different diseases. The data will enable us to determine, for example, the extent to which research related to a given disease tends to be concentrated in rapidly growing and advancing scientific fields (e.g., molecular genetics) as opposed to mature fields. They will also allow us to quantify the extent to which research on a disease utilizes innovative research techniques (e.g., protein engineering), and how much the distribution of techniques has changed over time.

At present, however, we must treat π_i as unobservable. If research productivity is uncorrelated across diseases with disease burden, that is, if differences in supply (or cost of achieving progress) are uncorrelated with differences in demand (or benefits of achieving progress), estimation of equation (5) will yield an unbiased estimate of the relationship between research expenditure and disease burden. It is possible, however, that N_i and π_i are negatively correlated: the diseases that impose the heaviest burden do so, in part, because of the low productivity of past research on those diseases (which should also have resulted in relatively low research funding on them). If this is the case, then the omission of π_i from the research expenditure equation would bias the estimated coefficient on $\ln N_i$ toward zero. In particular, although the theory implies that the coefficient on $\ln N_i$ should be greater than one, we should not be surprised if we obtain estimated coefficients smaller than one; in other words, if we fail to observe this kind of “increasing returns.”

In future research, I hope to directly estimate the contribution of medical research expenditure to subsequent progress against disease, by analyzing the correlation across diseases between research investment and indicators of progress, such as reductions in potential life years lost.¹¹ I recognize, however, that heterogeneous, unobserved research productivity is likely to lead to *overestimates* of the average return to research expenditure. Diseases receiving the greatest research funding are presumably those for which research productivity is highest. The slope of the relationship

respond to scientific advances, technology trends, and changes in the political and social environment.”

11. The existing evidence on the contribution of medical research expenditure to subsequent progress against disease is rather limited. The National Institutes of Health (1993) have produced estimates of cost savings from thirty-four “examples” of health care advances resulting from NIH support for applied research and clinical trials. Most focus on a single innovation such as a new vaccine, a new diagnostic test, or a particular therapy. But these case studies are not necessarily a random sample of all NIH-sponsored research, so they may not reveal the “aggregate or average” effect of this research on costs. It is possible, for example, that the distribution of cost savings is highly skewed to the right—a few programs confer large cost savings, but the majority confer few—and that the specific examples chosen tend to be concentrated in the upper tail of the distribution. Mushkin (1979) attempted to determine econometrically the contribution of biomedical research to reductions in mortality and morbidity. But most of her analysis was in an aggregate time-series framework and was based on fairly crude measures of biomedical research, such as the number of biomedical Ph.D.’s lagged ten years.

across diseases between research funding and progress exceeds the mean of the slopes of the disease-specific relationships.¹²

15.2 Data Sources and Methods

15.2.1 Data on Government-Funded Research Expenditures, by Disease

We have calculated distributions of government-funded biomedical research expenditure, by disease, from records of research grants contained in NIH's CRISP system. The CRISP database includes records of all research ventures supported by the U.S. Public Health Service since 1972. In fiscal year 1995, there were records of 63,289 grants, the total value of which was \$10.1 billion. Most of this research falls within the broad category of extramural projects: grants, contracts, and cooperative agreements conducted primarily by investigators at universities, hospitals, and other research institutions. The projects are funded by NIH and the Substance Abuse and Mental Health Services Administration. A very small number of these research grants are funded by the Centers for Disease Control, the Food and Drug Administration (FDA), the Health Resources and Services Administration, and the Agency for Health Care Policy and Research. CRISP also contains information on intramural research programs conducted by scientists employed by the FDA and the various institutes of the NIH.

Each record reports the name of the investigator, the name and address of his or her organization (e.g., university and department), the title (and in many cases an abstract) of the project, the administering organization (e.g., National Cancer Institute), the award amount (including both direct and indirect costs), the type of award, and a number of (generally about fifteen) indexing terms assigned by Technical Information Specialists in the Research Documentation Section, Information Systems Branch, of NIH's Division of Research Grants. The indexing process is governed by the CRISP thesaurus, which is the "controlled vocabulary used to assign indexing terms for the CRISP System, and to retrieve subject-related information from it."

The number of distinct indexing terms in the CRISP thesaurus is quite large (about nine thousand), but most of these terms are organized into

12. The reasoning underlying this is the same as that underlying Gary Chamberlain's (1984) argument that estimation of production functions using data for a cross-section of firms will result in overestimates of the returns to factors of production, such as labor. Firms with exogenously higher productivity (due, e.g., to greater managerial ability) will employ more workers. Chamberlain's point concerns the coefficient of one variable only—labor in a production function framework. If more than one variable is involved, their coefficients will not necessarily be biased toward zero: The direction of bias depends on the entire covariance matrix. Not all the coefficients on all the $\ln N_i$ variables will be biased toward zero.

a small number of hierarchical classification schemes, including one for diseases. Table 15.3 illustrates the disease classification; it is similar to the International Classification of Diseases, the system used for reporting diagnoses in most health-related data. There are thirty-five disease categories at the highest level of aggregation. Within each of these is a series of more specific disease categories. Space limitations prevent us from displaying the entire “tree structure” of diseases (which includes about twenty-nine hundred items), but to illustrate the classification system we show the second level classification of “nervous disorders” and a branch leading to

Table 15.3 **Classification Systems for Diseases Used in CRISP Database**

Blood disorder
Calcium disorder
Cardiovascular disorder
Communicable disease
Communication disorder
Congenital disorder
Connective tissue disorder
Digestive disorder
Ear disorder
Endocrine disorder
Enzyme deficiency
Eye disorder
Genetic disorder
Hernia
Immunopathology
Infection
Injury
Lymphatic disorder
Mental disorder
Metabolism disorder
Musculoskeletal disorder
Neoplasm/cancer
Nervous disorder
Autonomic disorder
Central nervous system disorder
Brain disorder
Cataplexy
Central nervous system neoplasm
Degenerative motor system disease
Encephalomyelitis
Gliosis
Hemiplegia
Meningitis
Infectious meningitis
Bacterial meningitis
Viral meningitis
Lymphocytic choriomeningitis
[Other disorders]

a “fifth level” disease (with no further subcategories), lymphocytic choriomeningitis.

This disease classification scheme enables us to compute distributions of research grants and dollars by disease, at various levels of aggregation.¹³ How accurate will these distributions be? Recently the Office of the Director of NIH prepared a report that included estimates of NIH fiscal year 1994 research support by disease. These figures, based on data provided by NIH institutes, centers, and divisions (ICDs), “reflect NIH-wide resources devoted to research on the listed diseases . . . [and] generally do not correspond to budget figures for the ICD identifying the cost data.”¹⁴ For sixteen randomly selected diseases, I compared fiscal year 1994 funding as reported there with the number of fiscal year 1995 grants citing the disease contained in the fiscal year 1995 CRISP database.

The raw data are reported in table 15.4. A scatter plot of the logarithms of these two variables is shown in figure 15.2; their correlation coefficient is .91. Despite differences in timing and unit of measurement, the two estimates of relative research support by disease are quite similar, suggesting that the CRISP data are reasonably reliable up to a first-order approximation.

As NIH officials observe, much NIH-sponsored research is basic in nature and, although “scientific advances would not have been possible without continuing insight and understanding regarding the fundamental mechanisms of life and disease . . . basic research linkages to health care advances are complicated, long-term, and impossible to allocate clearly” (NIH 1993, 3). Therefore, many research grants do not refer to *any* disease (even though the research may ultimately lead to breakthroughs in the treatment of that disease). In other words, the grants fall into two categories: those that have been assigned to at least one disease and those that have not been assigned.¹⁵ My estimates of research activity by disease are based only on grants that have been assigned.¹⁶ Due to the logarithmic specification of equation (6), the validity of my parameter estimates does not require me to reliably measure the *absolute level* of research funding, by disease; their validity is predicated only on reliable measurement of

13. Data on the disease distribution of *private* R&D sponsored by pharmaceutical firms are available from the Pharmaceutical Research and Manufacturers Association’s Annual Survey of companies. Unfortunately, the private R&D data are disaggregated into only about seven broad categories. Figure 15.1 shows the percentage distributions of both private and government R&D, by these categories. Public R&D seems to be more concentrated on digestive/genitourinary and neoplasm/endocrine/metabolic diseases, and less concentrated on infective/parasitic, nervous system, and cardiovascular diseases than private R&D.

14. NIH (1995), table 1.

15. This distinction resembles the distinction made in industrial R&D between basic and applied research.

16. When two or more diseases are cited by a grant, I assign the *entire* amount of funding for the grant to *each* of the diseases cited.

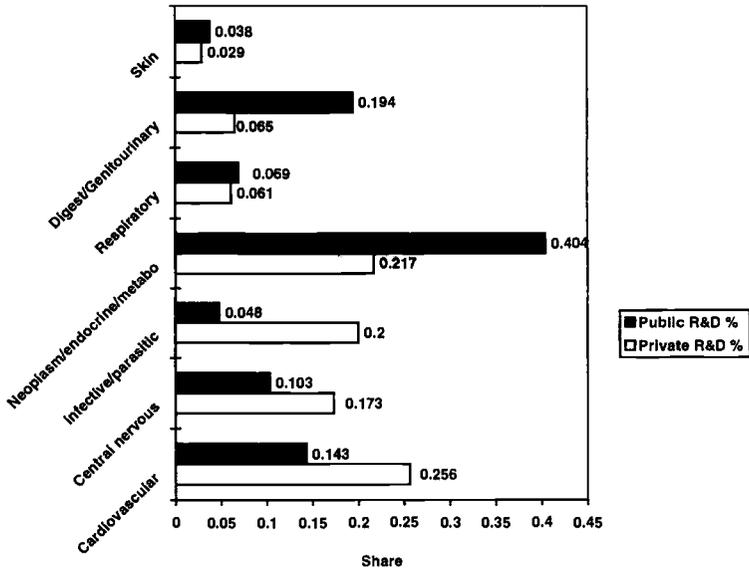


Fig. 15.1 Shares of public and private health R&D allocated to major disease categories in 1982

Sources: Public R&D: 1982 CRISP file; Private R&D: PhRMA Annual Survey.

Table 15.4 Comparison of Fiscal Year 1994 Funding with Fiscal Year 1995 Grants Citing the Disease, for Sixteen Randomly Selected Diseases

Disease/Disorder	FY 1995 Grants	FY 1994 Funds (million \$)
Diabetes	1,390	292
Epilepsy	338	52
Asthma	345	66
Arthritis	476	191
Atherosclerosis	650	116
Schizophrenia	458	111
Multiple sclerosis	123	78
Obesity	474	83
Osteoporosis	288	92
Parkinson's	253	68
Psoriasis	53	3
Sickle cell anemia	278	54
Suicide	94	17
Tuberculosis	248	50
Pneumonia and influenza	230	60

Source: NIH (1995), table 1; and CRISP database.

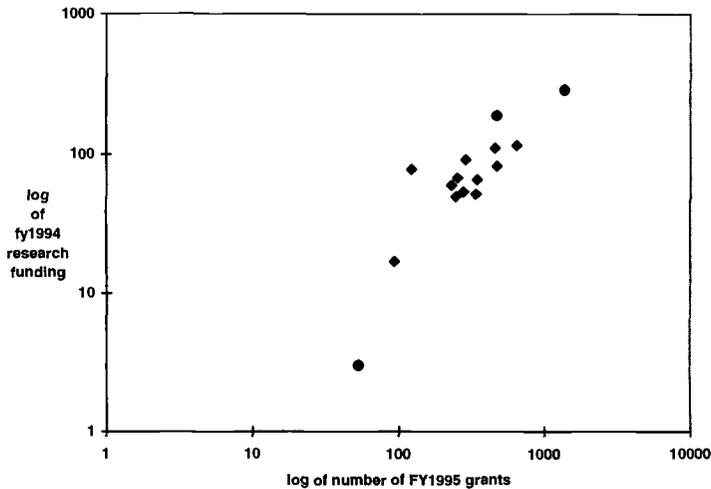


Fig. 15.2 Relationship between estimated NIH research funding, by disease, and number of NIH grants citing disease

relative research funding, or activity. Table 15.5 shows the fraction of 1972 and 1995 research grants whose indexing terms referred to any (at least one) disease and to specific diseases (at the highest level of aggregation) in the CRISP classification. In both years, about half of the grants referred to at least one disease.¹⁷ This is consistent with NIH's statement that "slightly over half, on average, of each Institute's budget supports the best research grant proposals regardless of specific applicability to prevention and treatment of a disease, but in expectation that their results will contribute to advances against diseases within their purview as well as diseases in other Institutes and to our knowledge generally." Relative emphasis on different diseases has been reasonably stable: the correlation across diseases (excluding pathology) between the 1972 and 1995 fractions is .85.

15.2.2 Data on Disease Burden, Prevalence, and Incidence

As indicated in equation (4) above, rather than treating disease burden N (or reduction in the quantity and quality of life) as a scalar, I regard it as an index of a number of disease mortality and morbidity attributes. Data on these attributes are obtained from two sources: the Vital Statistics–Mortality Detail file, a virtually complete census of deaths in the United States, and the National Health Interview Survey, a continuing nationwide survey of households for which a probability sample of the

17. The increase in this fraction, from 49 percent in 1972 to 57 percent in 1995, appears to be attributable to the large increase (from 9 to 24 percent) in the fraction of grants referring to "pathology."

Table 15.5 Percent of 1972 and 1995 NIH Grants Referring to Any Disease and to Specific Diseases

Disease (Ranked by % of 1995 Grants)	% of 1972 Grants	% of 1995 Grants
Any disease	48.5	56.7
Pathology	8.9	23.7
Neoplasm/cancer	8.6	12.2
Mental disorder	6.5	9.8
Nervous system disorder	6.2	9.6
Communicable disease	2.6	8.2
Immunopathology	4.4	7.8
Cardiovascular disorder	9.4	7.4
Metabolism disorder	6.7	5.4
Blood disorder	7.1	5.0
Digestive disorder	5.2	4.2
Respiratory disorder	3.5	4.0
Endocrine disorder	4.4	4.0
Reproductive system disorder	2.5	3.9
Infection	1.6	3.5
Lymphatic disorder	3.3	3.2
Congenital disorder	3.9	3.1
Musculoskeletal disorder	2.6	3.0
Injury	1.3	2.6
Urinary tract disorder	3.5	2.2
Eye disorder	1.9	2.0
Skin disorder	1.7	1.9
Genetic disorder	1.4	1.3
Communication disorder	1.0	1.2
Connective tissue disorder	0.9	0.7
Ear disorder	0.5	0.6
Pregnancy disorder	0.4	0.5
Nutrition disorder	2.2	0.5
Calcium disorder	0.8	0.3
Enzyme deficiency	0.0	0.2
Postnatal growth disorder	0.2	0.2
Syndrome	0.0	0.2
Orphan disease/drug	0.0	0.1
Nutrient intake disorder	0.1	0.1
Plant disease	0.0	0.0
Hernia	0.0	0.0

civilian noninstitutionalized¹⁸ population of the United States is interviewed by the U.S. Bureau of the Census regarding the health and other

18. It should be pointed out that the restriction of the NHIS to the civilian population not confined to institutions affects the estimated prevalence of chronic conditions. Omission of the institutionalized population reduces the prevalence estimates, especially for the elderly, because the proportion of persons in institutions who have chronic conditions is high. These estimates do not indicate the prevalence in the total population.

characteristics of each member of the household. (The sample for the years 1990–92 was composed of 142,638 households containing 368,075 persons.)

Our use of these two data sources reflects my belief that to obtain a reasonably complete accounting for disease burden, one must consider data on both the dying and the living. Analysis based on only one source will almost surely be subject to considerable sample selection bias.

Premature Mortality Data

The measure of disease burden I computed from the mortality file is potential life years lost before age sixty-five, by disease.¹⁹ The latter is defined as the summation of (65 – age-at-death) for decedents under sixty-five. This is a standard measure of disease burden, or (lack of) progress against disease, in health statistics. It has the drawback of giving no weight at all to deaths of people aged sixty-five and over.

Data on Prevalence of Selected Chronic Conditions

Collins (1997) presents statistics on the prevalence of selected chronic conditions in the United States during 1990–92 by age, sex, race, family income, and geographic region, derived from data collected in the National Health Interview Survey (NHIS). He also reports the percent of selected conditions that cause activity limitation, the percent for which a physician was consulted, and the percent that caused hospitalization.

All information collected during the survey is from responsible family members residing in the household. Methodological studies have shown that chronic conditions are generally underreported in interview surveys. Respondents in health interviews tend to report conditions of which they are aware and about which they are willing to report to the interviewer. Reporting is better for conditions that have made a significant impact on affected individuals and their families. Conditions that are severe or costly, or are being treated, tend to be better reported than conditions having less impact. Methodological studies have also indicated that inclusion of a checklist of descriptive condition titles as part of the interview questionnaire increases the probability that a respondent will recognize the terms and report those of which the respondent is aware.

The current procedure for collecting information on chronic conditions was established in 1978. Currently, six categorical lists of selected chronic conditions are included in the NHIS questionnaire: circulatory conditions; respiratory conditions; digestive conditions; impairments and conditions of the nervous system and sense organs; conditions of the skin and subcu-

19. Demographic information on the death certificate is provided by the funeral director based on information supplied by an informant. Medical certification of cause of death is provided by a physician, medical examiner, or coroner.

taneous tissue and of the musculoskeletal system and connective tissue; and endocrine, nutritional, and metabolic diseases and immunity disorders, diseases of the blood and blood-forming organs, and conditions of the genitourinary system. Each family in the NHIS is questioned on only one of these six lists, selected on a predetermined basis. Therefore, each list is administered to only one-sixth of the total NHIS sample each year. For some items, responses are based on the following question: "During the past 12 months did anyone in the family (read names) have . . . ?" For others, responses are based on the question "Does anyone in the family (read names) now have . . . ?" For the rest, responses are based on the question "Has anyone in the family (read names) ever had . . . ?" Estimates for days of disability caused by chronic conditions are based on the number of disability-days reported for the two weeks before interview.

The survey includes data only on persons living in the household at the time of interview. Thus the experience of persons who died prior to the time of interview is excluded from the data. Also excluded is the experience of persons who were institutionalized or who were members of the armed forces at the time of the household interview.

In these data, "prevalence" is defined as the average number of some item existing during a specified interval of time—usually referred to as "period prevalence"—rather than the number of some item existing at a given point in time—usually referred to as "point prevalence." Chronic conditions are defined as conditions that either were first noticed three months or more before the date of interviews, or belong to a group of conditions considered chronic regardless of when they began.

The data presented represent the prevalence of conditions, not the prevalence of persons with a chronic condition. However, for most conditions, the condition prevalence and the person prevalence are almost identical.²⁰

15.3 Preliminary Estimates

15.3.1 Premature Mortality

The first measure of disease burden I analyze is potential life years lost before age sixty-five (PLYL). Data on PLYL in 1980 and government re-

20. There are some instances in which large variations are present; these occur for two different reasons. The first is that a prevalence estimate of a condition may include more than one of the specified checklist items or a checklist item and a specified "other condition" item that falls into the same International Classification of Diseases category as the checklist item. The second reason is that some prevalence categories shown are a combination of other categories and, as a result, a person may have more than one of the conditions that are added to form the combined category. The concept of condition prevalence is generally used in NHIS because specific health indexes such as limitation of activity and disability days can be ascribed to specific conditions. In addition, prosthetic and pharmaceutical treatment modes are more condition specific than person specific.

Table 15.6 Life Years Lost before Age Sixty-Five in 1980, and Public R&D Expenditure in 1982, Fourteen Major Disease Categories

Disease/Disorder ^a	Life Years Lost before Age Sixty-Five in 1980	Public R&D Expenditure in 1982 (million \$)
Diseases of the circulatory system (390–459)	2,043,559	117
Neoplasms (140–239)	1,860,531	113
Congenital anomalies (740–759)	760,820	37
Diseases of the digestive system (520–579)	503,531	27
Diseases of the respiratory system (460–519)	434,770	47
Diseases of the nervous system and sense organs (320–389)	294,239	118
Endocrine, nutritional, and metabolic diseases and immunity disorders	223,015	168
Infectious and parasitic diseases (001–139)	162,568	33
Mental disorders (290–319)	124,407	70
Diseases of the genitourinary system (580–629)	86,015	73
Diseases of the blood and blood-forming organs (280–289)	49,814	16
Diseases of the musculoskeletal system and connective tissue	37,403	25
Complications of pregnancy, childbirth, and the puerperium	12,536	8
Diseases of the skin and subcutaneous tissue (680–709)	7,848	21

^aNumbers in parentheses are International Classification of Diseases codes.

search funding, in 1982, for fourteen major disease categories are shown, in descending PLYL order, in table 15.6. Diseases of the circulatory system and neoplasms are, by far, the diseases with the largest tolls in terms of premature death. While the research funding for these two diseases is among the highest for all diseases, R&D funding for two other diseases with much smaller burdens exceeds the funding for the first two diseases, in one case by a large amount. Nevertheless, as the scatter plot in figure 15.3 and the following regression indicate, there is a very strong positive relationship across the entire sample between life years lost and public R&D expenditure (*t*-statistics in parentheses):

$$\ln(\text{RD82}) = -0.464 + 0.355 \ln(\text{LYL80}) + e \quad R^2 = .459$$

$$(0.34) \quad (3.19) \quad N = 14.$$

Life years lost in 1980 explains almost half of the variation across diseases in 1982 research expenditure. However, contrary to the implication of my simple theoretical model of research allocation, the coefficient on $\ln(\text{LYL80})$ is significantly less than one. As argued above, this may be due

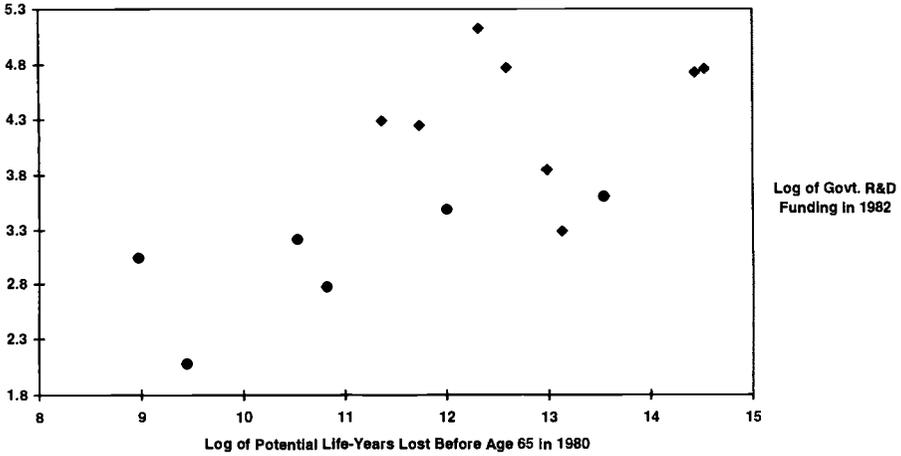


Fig. 15.3 Relationship between government research funding, by disease, in 1982, and life years lost before age sixty-five, by disease, in 1980

to a negative correlation between the regressor and the omitted research-productivity variable.²¹

Life years lost can be classified by sex, race, educational attainment, and other characteristics, so we can investigate whether premature mortality among certain demographic groups tends to be associated with especially high government research funding. Sixty percent of life years lost before age sixty-five are lost by males, and 25 percent are lost by nonwhites (who make up about 10 percent of the population), reflecting the lower life expectancy of these two groups. The proportion of life years lost by men and by nonwhites varies considerably across diseases. Whites account for 81 percent of life years lost to neoplasms but for only 53 percent of those due to diseases of the blood and blood-forming organs. Men account for 81 percent of life years lost to infectious and parasitic diseases but

21. I obtain quite similar results when I use data covering other time periods or when I substitute life years lost before age eighty for life years lost before age sixty-five. (About twice as many life years are lost before age eighty as are lost before age sixty-five; the correlation across diseases between the two is very high—.98.) The correlation coefficient between $\ln(\text{RD}82)$ and $\ln(\text{LYL}80)$ is .677. The correlation coefficients between the log of the number of year t ($t = 1980, 1995$) NIH grants referring to disease i and the log of life years lost before age j ($j = 65, 80$) to disease i in year t are as follows:

	$t = 1980$	$t = 1995$
$j = \text{age } 80$.764	.710
$j = \text{age } 65$.739	.677

for only 28 percent of life years lost to musculoskeletal and connective-tissue diseases.

The matrix of correlation coefficients for four variables— $\ln(\text{RD82})$, $\ln(\text{LYL80})$, and the fractions of life years lost to men (%MALE) and to whites (%WHITE)—are reported in table 15.7. Public R&D investment is significantly positively correlated with the fractions of life years lost to men and (especially) to whites, as well as with the total number of life years lost. Indeed, %WHITE is more strongly correlated with R&D than total life years lost is. (A scatter plot of $\ln[\text{RD82}]$ against %WHITE is shown in fig. 15.4.) But as the second column of coefficients reveals, both %MALE and %WHITE are significantly positively correlated with total life years lost: the diseases associated with the greatest number of premature deaths are those for which men and whites account for the greatest

Table 15.7 Correlation Matrix for Four Variables

	$\ln(\text{RD82})$	$\ln(\text{LYL80})$	%MALE
$\ln(\text{LYL80})$	0.67714 (0.0078)		
%MALE	0.55375 (0.0399)	0.56093 (0.0369)	
%WHITE	0.75643 (0.0017)	0.88477 (0.0001)	0.50235 (0.0672)

Note: Numbers in parentheses are *p*-values.

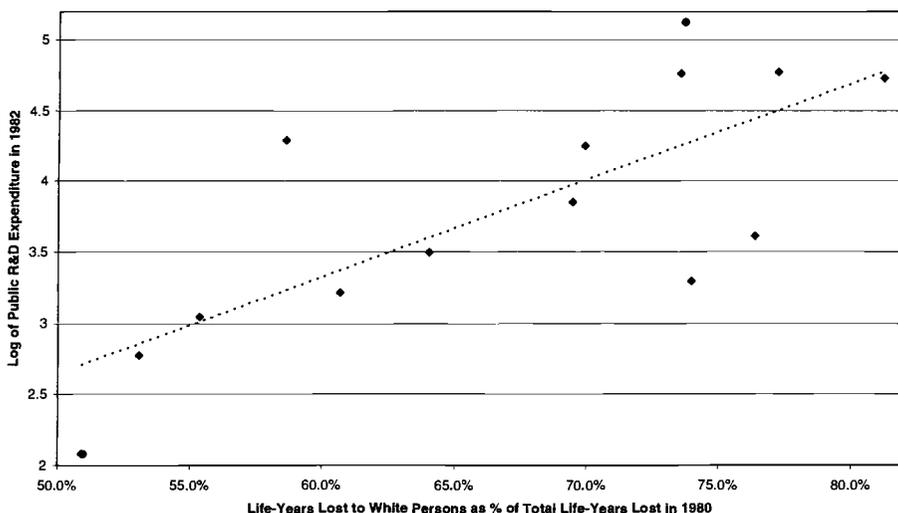


Fig. 15.4 Relationship across diseases between public R&D expenditure and percent of life years lost to white persons

fractions of life years lost. We therefore need to determine whether %MALE and %WHITE have significant effects on public R&D, controlling for total life years lost (although our ability to determine this will be hampered by multicollinearity). The appropriate regressions are

$$\ln(\text{RD82}) = -0.164 + 0.280\ln(\text{LYL80}) + 1.11\% \text{MALE} + e \quad R^2 = .503 \\ (0.12) \quad (2.08) \quad (0.99) \quad N = 14,$$

$$\ln(\text{RD82}) = -0.813 + 0.019\ln(\text{LYL80}) + 6.56\% \text{WHITE} + e \quad R^2 = .573 \\ (0.64) \quad (0.09) \quad (1.71) \quad N = 14.$$

The coefficient on %MALE is insignificant and the inclusion of this variable only slightly reduces the coefficient on $\ln(\text{LYL80})$. In contrast, the coefficient on %WHITE is marginally significant, even in the presence of the other regressor, which becomes insignificant (with a t -statistic of only 0.09) when %WHITE is included. We also estimated an alternative functional form of the relationship $\text{RD82} = f(\text{LYL80}, \% \text{WHITE})$:

$$\ln(\text{RD82}) = 2.29 + 1.35\ln(\text{LYL80} \times \% \text{WHITE}) \\ (1.14) \quad (1.90) \\ -1.30\ln[\text{LYL80} \times (1 - \% \text{WHITE})] + e \quad R^2 = .560, \\ (1.44) \quad N = 14.$$

These estimates indicate that research expenditure is positively correlated with life years lost by whites but not by nonwhites; the coefficient on the latter is negative, but its p -value is only .18. The two coefficients are virtually equal in magnitude and opposite in sign; if one imposes that restriction (which is not rejected by the data), the estimates are

$$\ln(\text{RD82}) = 2.72 + 1.47\ln[\% \text{WHITE}/(1 - \% \text{WHITE})] + e \quad R^2 = .558 \\ (8.40) \quad (3.89) \quad N = 14.$$

The data are highly consistent with the hypothesis that the amount of publicly funded research on a disease decreases with the share of life years before age sixty-five lost to the disease that are lost by nonwhites. A possible explanation for this finding is that lack of scientific knowledge is a less important cause of premature mortality among nonwhites than it is among whites. Nonwhite premature mortality may be due, to a greater extent, to poor diet, reduced utilization of medical care, or other factors. In other words, it is plausible that the health status of nonwhites tends to be well below the frontier of medical knowledge, whereas the health status of whites tends to be on, or closer to, the frontier. The purpose of biomedical research is to shift the frontier outward, and the allocation or "direction" of research should depend (more) on the distribution of the disease burden of those on, or close to, the frontier. If cures for diseases that im-

pose a heavy toll on minorities have already been found, then the productivity of further research on those diseases may be quite low.

The relative lack of research on diseases borne disproportionately by minorities may also be due to other reasons and may not be efficient. It may reflect the relatively low representation of minorities among the ranks of biomedical scientists. The National Science Foundation monitors the participation of women and minorities in science and engineering and has adopted some policies to increase their participation.

15.3.2 Prevalence and Severity of Chronic Conditions in the (Living) Population

Table 15.8 presents data on the number of FY1995 research grants mentioning chronic conditions surveyed in the National Health Interview Survey and the number of people having, and limited in activity by, these conditions.²² The condition mentioned in the most (1,807) research grants is diabetes. About seven million Americans suffer from diabetes, according to this household survey; about one-third of them are limited in activity by this condition. Although arthritis is far more prevalent, afflicting thirty-two million Americans, the number of research grants mentioning it (609) is much smaller.

Table 15.9 presents correlation coefficients of the logarithms of these variables and related measures of condition severity. This table indicates that the number of research grants mentioning a chronic condition has a very small and insignificant correlation with the number of people with the condition and with the number who have seen a physician about that condition. Research activity is weakly positively related (p -value = .08) to the number of people who have been hospitalized for a condition. It is very strongly positively related (p -value = .0003) to the number of people whose activities are limited by that condition. Somewhat surprisingly, research activity is significantly positively correlated with the *proportion* of people who have seen a doctor or been hospitalized, as well as those whose activities are limited.²³

The determinants of the number of research grants citing chronic conditions are further analyzed in table 15.10. The first column presents the regression of $\ln(\text{NGRANTS95})$ on a measure of condition prevalence

22. In this section the measure of public research activity I use is the number of grants rather than the dollar value of those grants. For technical reasons, the former is much easier to compute. Substitution of the former for the latter will not affect my results if the average size of grants is uncorrelated across conditions with the number of grants. In the future I plan to compute the distribution of dollars by condition and to integrate the premature mortality and chronic-condition prevalence analyses.

23. This is particularly surprising since, as the second column of table 15.9 indicates, these proportions are significantly inversely related to prevalence per se: conditions that are more prevalent tend to be less severe (i.e., associated with lower probabilities of hospitalization, activity limitation, and physician consultation).

Table 15.8 **Number of FY1995 Research Grants Citing, and Number of People Reporting and Limited in Activity by, Major Chronic Conditions**

NGRANT	N	NLA	Chronic Condition
1,807	6,962	2,416	Diabetes
1,540	27,600	2,926	High blood pressure (hypertension)
671	1,293	374	Diseases of retina
609	31,788	6,739	Arthritis
593	1,513	79	Diseases of prostate
573	3,739	157	Anemias
493	11,482	2,503	Asthma
425	1,243	552	Epilepsy
402	1,562	1,367	Mental retardation
315	8,169	1,291	Blindness and other visual impairments
293	766	130	Liver diseases including cirrhosis
288	3,002	1,078	Cerebrovascular disease
282	23,266	1,280	Deafness and other hearing impairments
258	7,732	2,436	Ischemic heart disease
241	180	125	Multiple sclerosis
241	2,725	556	Speech impairments
216	802	190	Malignant neoplasm of breast
203	6,416	391	Cataracts
195	2,433	326	Glaucoma
119	2,333	161	Enteritis and colitis
118	741	133	Congenital heart disease
118	834	45	Disease of the esophagus
103	322	200	Malignant neoplasms of stomach intestines
99	218	132	Malignant neoplasms of lung bronchus
90	1,984	18	Menstrual disorders
85	2,378	50	Psoriasis
84	1,861	821	Emphysema
84	1,325	46	Kidney infections
84	1,911	134	Tachycardia or rapid heart
83	7,868	504	Heart rhythm disorders
83	344	76	Malignant neoplasm of prostate
80	4,201	328	Ulcer gastric duodenal and/or peptic
74	2,269	88	Malignant neoplasms of the skin
61	2,074	199	Hardening of arteries
59	3,121	240	Gastric ulcer
54	73	2	Benign neoplasm of breast
53	217	7	Cleft palate
50	3,003	30	Gastritis and duodenitis

Notes: NGRANT = Number of FY1995 grants mentioning condition. N = Average number of people (in thousands) in 1990–92 reporting that they have the condition. NLA = Average number of people (in thousands) in 1990–92 reporting that their activities are limited by the condition. Only conditions cited by fifty or more grants are listed.

Table 15.9 Correlations between Research Activity and Prevalence/Severity of Chronic Conditions

	LGRANTS	LN	LNLA	LNHOSP
LGRANTS: log(no. of research grants)	1.00 (0.00)			
LN: log(no. of people w. condition)	0.04 (0.74)	1.00 (0.00)		
LNLA: log(no. w. limited activity)	0.40 (0.00)	0.54 (0.00)	1.00 (0.00)	
LNHOSP: log(no. hospitalized)	0.20 (0.08)	0.61 (0.00)	0.74 (0.00)	1.00 (0.00)
LNPHYS: log(no. seeing physician)	0.07 (0.53)	0.99 (0.00)	0.59 (0.00)	0.66 (0.00)
LA: % w. limited activity	0.35 (0.00)	-0.26 (0.00)	0.49 (0.00)	0.12 (0.19)
HOSP: % hospitalized	0.22 (0.06)	-0.46 (0.00)	0.09 (0.31)	0.28 (0.00)
PHYS: % seeing physician	0.32 (0.00)	-0.43 (0.00)	0.18 (0.05)	0.18 (0.04)

Note: Figures in parentheses are probability values.

Table 15.10 Determinants of Number of FY1995 Research Grants Mentioning Chronic Conditions ($N = 54$)

	Equation 1	Equation 2	Equation 3
ln(N)	0.142 (0.73)		
%LA	4.45 (2.77)		
ln($N \times \%LA$)		0.651 (4.12)	0.369 (2.39)
ln[$N \times (1 - \%LA)$]		-0.436 (2.17)	-0.167 (0.88)
%INCOME < \$10,000			8.61 (2.17)
%AGE < 18			5.67 (2.63)
%AGE > 75			7.30 (2.73)
Intercept	2.09 (1.31)	3.82 (2.81)	0.267 (0.17)
R^2	0.1317	0.2460	0.4470

Notes: The dependent variable is the log of the number of FY1995 grants. N = Average number of people (in thousands) in 1990-92 reporting that they have the condition. %LA = Fraction of people reporting that their activities are limited by the condition. %INCOME < \$10,000 = Fraction of people with household income < \$10,000. %AGE < 18 = Fraction of people under eighteen years of age. %AGE > 75 = Fraction of people over seventy-five years of age. Numbers in parentheses are t -statistics.

($\ln[N]$) and severity (%LA). As one might expect given the simple correlations in the previous table, only the severity measure has a significant positive effect on research activity. In the second column, I estimate an alternative functional form of the relationship; the regressors are the logarithms of the number of people with the condition whose activities are ($N \times \%LA$) and are not ($N \times [1 - \%LA]$) limited by the condition. The coefficient on the former is positive and highly significant, indicating that the amount of public research about a chronic condition increases with the number of people whose activities are limited by that condition.²⁴ Moreover, the amount of public research is significantly *inversely* related to the number of people who have a condition but whose activities are not limited by it. This could conceivably signify that, the greater the number of people who have a condition but are not seriously affected by it, the greater the odds that an adequate treatment for the condition already exists, and the less worthy that condition is of further research. This inverse relation becomes insignificant, however, when we include (in column 3) measures of the income and age distribution of persons reporting the condition. This regression indicates that there tends to be more research about chronic conditions that are prevalent among people living in low-income (below \$10,000) households, and that are prevalent among the young (under age eighteen) and the old (above age seventy-five). This suggests that the poor, the young, and the very old may derive disproportionately large benefits from government-sponsored biomedical research. In the previous section I reported that the amount of publicly funded research on a disease decreases with the share of life years before age sixty-five lost to the disease that are lost by nonwhites. Since nonwhites are more likely to be poor than whites, it is surprising that chronic conditions prevalent among the poor tend to be more intensively researched.

15.4 Summary

I have developed a simple theoretical model of the allocation of the applied component of public biomedical research expenditure—the approximately 50 percent that is of direct, near-term relevance to specific diseases—and presented some empirical evidence about the determinants of this allocation. The implications of the theoretical model are consistent with government officials' descriptions of the allocation process: the structure of expenditure should depend upon research productivity (or “scientific opportunity”) as well as on public health need, or the societal and economic burden of the disease/condition.

Although we lack, at this point, useful indicators of research productiv-

24. As in the analysis of premature mortality, however, the elasticity is significantly less than unity.

ity (i.e., of the *cost* of achieving research advances), we have a number of measures of disease burden (i.e., of the potential *benefit* of achieving these advances). Analysts of technological change typically have data on neither the costs nor the benefits of technical advance. Failure to measure research productivity will not necessarily bias my estimates; if it does, it seems likely to bias them toward zero.

I calculated distributions of government-funded biomedical research expenditure, by disease, from records of all research projects supported by the U.S. Public Health Service; in fiscal year 1995, there were records of 63,289 projects whose total value was \$10.1 billion. Some research expenditure cannot be assigned to specific diseases, in some cases because the research being conducted is basic in nature. The distribution of research expenditure by disease that I constructed is quite similar to one calculated by NIH based on data provided by NIH institutes, centers, and divisions (ICDs) designed to "reflect NIH-wide resources devoted to research on the listed diseases" (as opposed to budget figures for the ICD identifying the cost data).

I performed an empirical examination of the relationship of public research expenditure to a number of measures of disease burden. To avoid "sample selection bias," and to obtain a reasonably complete accounting of disease burden, I utilized data on both the dying (from the Vital Statistics–Mortality Detail file) and the living (from the National Health Interview Survey).

The mortality-related measure of disease burden I use is life years lost before age sixty-five. I found a very strong positive relationship across diseases between total life years lost and public R&D expenditure (although the slope of this relationship was smaller than that implied by the theory, perhaps due to failure to measure research productivity). Further analysis indicated that research expenditure is positively correlated with life years lost by whites but not with life years lost by nonwhites. In other words, the amount of publicly funded research on a disease decreases with the share of life years before age sixty-five lost to the disease that are lost by nonwhites. A possible explanation for this finding is that lack of scientific knowledge is a less important cause of premature mortality among nonwhites than it is among whites.

Disease prevalence and severity data for the (living) population provide additional indicators of disease burden. I found that the number of research grants mentioning a chronic condition has a very low and insignificant correlation with the number of people with the condition and with the number who have seen a physician about that condition. Research activity is weakly positively related to the number of people who have been hospitalized for a condition, and very strongly positively related to the number of people whose activities are limited by that condition. Moreover, there tends to be more research about chronic conditions that are prevalent

among people living in low-income households, and that are prevalent among the young (under age eighteen) and the old (above age seventy-five).

References

- Adams, James D. 1990. Fundamental stocks of knowledge and productivity growth. *Journal of Political Economy* 98 (4): 673–702.
- Chamberlain, Gary. 1984. Panel data. In *Handbook of econometrics*, vol. 2, ed. Z. Griliches and M. D. Intriligator, chap. 22. New York: Elsevier Science.
- Collins, J. G. 1997. Prevalence of selected chronic conditions: United States, 1990–1992. *Vital Health Statistics* 10 (194): 1–89.
- Cutler, David. 1995. Technology, health costs, and the NIH. Unpublished paper prepared for NIH Roundtable on Economics, 19 October.
- Garber, Alan, and Paul Romer. 1995. Evaluating the federal role in financing health-related research. Unpublished paper prepared for NIH Roundtable on Economics, 19 October.
- Henderson, Rebecca, and Iain Cockburn. 1996. Scale, scope, and spillovers: The determinants of research productivity in drug discovery. *RAND Journal of Economics* 27 (1): 32–59.
- Institute of Medicine. Committee on the NIH Research Priority-Setting Process. 1998. Scientific opportunities and public needs: Improving priority setting and public input at the National Institute of Health. Washington, D.C.: National Academy Press. Available at: <http://www.nap.edu>.
- Lichtenberg, Frank. 1996. Do (more and better) drugs keep people out of hospitals? *American Economic Review* 86:384–88.
- . 2000. The effect of pharmaceutical utilisation and innovation on hospitalisation and mortality. In *Productivity, technology, and economic growth*, ed. B. van Ark, S. K. Kuipers, and G. Kuper. Boston: Kluwer Academic.
- Mushkin, Selma. 1979. *Biomedical research: Costs and benefits*. Cambridge, Mass.: Ballinger.
- National Center for Health Statistics. 1995. *Health, United States, 1994*. Hyattsville, Md.: Public Health Service.
- National Institutes of Health. 1993. *Cost savings resulting from NIH research support*, 2nd ed. NIH Publication no. 93-3109, September. Washington, D.C.: National Institutes of Health.
- . 1995. Disease-specific estimates of direct and indirect costs of illness and NIH support. Washington, D.C.: National Institutes of Health.
- . 1997. Setting research priorities at the National Institutes of Health. Working Group on Priority Setting, September. Available at <http://www.nih.gov>.
- . N.d. Stories of discovery: NIH's contributions to progress against disease. Unpublished paper.
- Toole, Andrew. 2000. The impact of public basic research on industrial innovation: Evidence from the pharmaceutical industry. Policy paper no. 00-07. Institute for Economic Policy Research, Stanford University, November.
- Viscusi, W. Kip. 1995. Valuing the health consequences of biomedical research. Unpub. paper prepared for NIH Roundtable on Economics, 19 October.

This Page Intentionally Left Blank