Price Indexes for Clinical Trial Research: A Feasibility Study

Ernst R. Berndt
Iain M. Cockburn

December 26, 2012

Research support from the National Science Foundation, Collaborative Research: "An Experimental Producer Price Index for Clinical Trials" to the National Bureau of Economic Research, Award Number 0915674, and Boston University, Award Number 0915677, is gratefully acknowledged. We thank MediData Solutions Worldwide, Inc. ("MediData") for supplying database extracts, Ed Seguine (formerly at Fast Track Solutions, Inc. and MediData) for helpful discussions, and Rafael Campo and other individuals at MediData for helpful comments and data support. Rania Gihleb provided able research assistance. The opinions expressed herein are those of the authors, and do not necessarily reflect those of the National Science Foundation and MediData.

Document Name: Berndt-Cockburn PPI for RCTs V9.docx Date: 26 Dec 2012

DRAFT

I. Introduction

Clinical trials involving human subjects are a significant component of R&D expenditures in the US economy, with important implications for human health, physician practice and revenues, national and regional economic performance, and academic medicine. Of the $46.4 billion spent by Pharmaceutical Research and Manufacturers of America (PhRMA) member companies in 2010 on R&D, $32.5 billion (70%) was spent on Phase 1through Phase 4 clinical trials.[1] In recent years, clinical research has accounted for about 1/3 of the total NIH budget ($10.7 billion out of $30 billion in FY2010) of which a substantial fraction ($3.2 billion in FY2010) is expenditure on clinical trials.[2] . But while the expenditures associated with this aspect of development of new drugs continue to grow very rapidly, by some indicators the pace of innovation has slowed. For example while total R&D spending by PhRMA member companies has almost doubled over the last decade[3] (as has the overall NIH budget[4]), the number of new drugs and biologics approved by the FDA each year in the last decade has, at best, been static, and considerably less than during the 1990s. One prominent study (DiMasi et. al [2003]) reports that the capitalized cost of bringing a new drug to market, adjusted for general inflation in year 2000 dollars, more than doubled from $318 to $802 million between 1991 and 2003.

This raises some very basic—and as yet unanswered—questions. Increases in the cost per approved drug are often equated with "the price of innovation", but in fact little is known about how much of the increase in expenditure reflects changes in the prices of inputs to biomedical research and how much reflects changes in the quantity of research being performed. Have the prices of inputs to clinical research increased more rapidly than overall inflation, or are these inputs being used more intensively, or are both occurring? Moreover, to what extent has the "quality" of inputs changed? The growing complexity of clinical trials and the underlying science suggests that more time, more highly trained personnel, and more sophisticated equipment may be required to conduct a typical study.[5]

Very little data is currently available to inform discussion of such issues. While data are captured for some inputs to clinical research, such as salaries of post-doctoral fellows, relatively little is known about other important inputs to clinical research such as site administration costs, computational time, materials and investigator salaries. Critically, even where good data are

available on input prices, it is important to take into account how inputs are combined by focusing on an appropriate unit of analysis.

The highly influential studies by Cutler and coauthors on the costs of treating heart attacks have had a major impact on analyses of health expenditures by focusing attention on changes in the cost of an "episode of care" due to input substitution, rather than on changes in the per unit-price of inputs to care.[6] This observation may be particularly relevant in analyzing biomedical clinical research since this activity is widely believed to have become more costly not just because of increases in input costs, but because it has become more complex, more time-consuming, and more resource-intensive. Any effort to understand the causes and consequences of rising expenditure on clinical research must take these changes into account. The importance of such detailed research has recently been emphasized by a Conference Board/NSF workshop, which concluded, *inter alia*, that in order to support a new micro-to-macro research data infrastructure, "…comprehensive data on innovation input costs could be collected according to concepts used in modern business organizations."[7]

More generally, in the language of the economics of price measurement, we need to think about "constant-quality" changes in prices and quantities, i.e., hold the characteristics of the input and output activities constant when looking at changes in expenditures over time or cross-sectionally. Failure to do so can result in quite misleading interpretations and policy recommendations. Analyses of expenditures on computers, for example, recognize that there have been huge increases in the performance or capacity of the products sold, but very small changes in their nominal prices; "constant quality" prices have thus fallen substantially over time—various estimates suggest sustained real price declines of more than 25% per year over several decades.[8] Various governmental statistical agencies now routinely take this phenomenon into account for many types of information technology and other electronic goods in developing estimates of GDP, with quite marked impacts on measures of economic growth and productivity.[9]

While it is important, therefore, to quantify the "price" versus "quantity" component changes in R&D, adjusting both for quality, characterizing scientific research presents some very substantial measurement problems. Research activities are typically highly heterogeneous and idiosyncratic in nature, drawing on quite different inputs and resources to produce "output" which is very difficult to measure consistently. However in one respect, clinical trials may be

unusually tractable. Clinical development is a highly structured activity, in which individual "experiments" are relatively well-defined and activity is closely tracked. Industry trends are also creating an unusual opportunity to investigate these questions. While biopharmaceutical companies and non-profit entities continue to be the lead sponsors of clinical trials, much of the effort in conducting them is increasingly outsourced to contract research organizations ("CROs") rather than being incurred "in house". At least within the US, the investigators who recruit, treat, and observe subjects are drawn less from academic medical centers and increasingly more from independent physician practices.[10] This has meant that data on contractual terms among all these parties are now ever more important and increasingly visible.

The only existing R&D price index of which we are aware is the Biomedical Research and Development Price Index (BRDPI), published by the National Institutes of Health under agreement with the U.S. Bureau of Economic Analysis. The BRDPI measures changes in the weighted-average of the prices of all the inputs (e.g., personnel services, various supplies, and equipment) that are purchased or leased with the NIH budget to support research. Input weights reflect the changing actual shares of total NIH expenditures on each of the types of inputs purchased. According to the NIH,

> "Theoretically, the annual change in the BRDPI indicates how much NIH expenditures would need to increase, without regard to efficiency gains or changes in government priorities, to compensate for the average increase in prices due to inflation and to maintain NIH-funded research activity at the previous year's level."[11]

The BRDPI is published annually on a federal government fiscal year (October 1 – September 30) basis.

In this manuscript we report results from analyses of a sample of over 215,000 contracts regarding payments made by trial sponsors (directly or through CRO intermediaries) to clinical investigators and study sites from the PICAS® database maintained by MediData Worldwide Solutions, Inc. This sample covers over 24,000 distinct Phase 1 through Phase 4 clinical study protocols conducted between 1989 and 2011 in 52 different countries. Using information on the protocol characteristics we compute hedonic price indexes that allow us to estimate the rate of inflation in this particular aspect of clinical research, controlling for changes in the characteristics of clinical trials over the sample period. We find that while our measure of unit costs of this aspect of conducting clinical trials rose rapidly over the two decades covered by this

sample at 8% per year (roughly twice the rate of inflation in the NIH's Biomedical R&D Price Index), these changes in nominal costs appear to be driven by a variety of factors other than input costs. At least in this sample there has been a substantial increase in the level of effort required by investigators, and significant changes in both the composition of the sample across therapeutic classes and stages of clinical development, as well as in the organization of trials with a trend towards smaller numbers of patients per site and considerable variation over time in the geographic distribution of ex-US sites. After controlling for these factors using hedonic regression methods, we find much lower growth rates in costs, with adjusted rates of inflation between 1/3 and 2/3 lower than those seen in the unadjusted data.

II. Data

With the co-operation of MediData Worldwide Solutions, Inc. ("MediData") we assembled a dataset of more 216,076 observations on "investigator grants," which are payments made by a trial sponsor to the individual investigators or "sites" that enroll subjects.[12] These payments cover the investigators' costs of recruiting subjects, administering the treatment, measuring clinical endpoints, etc., plus overhead allowances reflecting payments for the use of the site's facilities. We focus on the total grant cost per patient ("TGPP") as the economically meaningful unit of analysis for understanding price trends. Total grant cost per patient is the total amount paid by the sponsor under its contract with the site, divided by the number of patients planned to be enrolled at that site. For about 12% of the records contained in the PICAS® database, the contract specifies only a per-patient amount, not the number of patients. These contracts are excluded from the results reported below, since we are unable to control for the scale of the site's effort. For ex-US sites where the contract is in a foreign currency, we convert to US dollars using the spot exchange rate. Typically these payments make up about half of the total cost of a trial, the remainder being "overhead" in the form of data management, site selection and monitoring, etc. by the sponsors.

Table I shows a summary of the number of records in the dataset by year each investigator contract was signed, along with summary statistics for the total grant cost per patient. The dataset used here covers the almost quarter century period 1989 to 2011. The number of records per year varies over time, with the period 1992-2002 accounting for almost 75% of the total number of records. The number of records per year in our sample reached a

peak in 2000 and declines steadily thereafter, reflecting two factors. First, the PICAS® database was originally compiled from an archive of paper records, and has since transitioned to electronic source documents. This transition led to a temporary decline in the number of contributions from the participating organizations in 2004-2005. Second, the fraction of contracts that do not specify the number of patients expected to be enrolled at the site (and are excluded from our sample) has increased over the past decade. This trend likely reflects tighter "real time" tracking and control of patient enrollment by trial sponsors.

As can be seen from the table, the mean TGPP rises quite rapidly in nominal terms, just over four-fold over the period 1989 to 2011, from $3773 to $16567, with an average annual growth rate (AAGR) of 7.5%;[13] the median value of each year increases slightly more rapidly, from $2779 to $13222, an AAGR of 8.2%.[14] By comparison, between fiscal years 1989 and 2011 the NIH's BRDPI increased much more slowly, barely doubling at an AAGR only half as large at 3.7%.[15] The distribution of TGPP is quite skewed, with the median somewhat below the mean value; a visual plot suggests that TGPP can be reasonably approximated with the lognormal distribution. Notably, the within-year coefficient of variation is relatively large but stable at around 0.80 at both the beginning and end of the sample period. While we attempt to account for this variation in TGPP with measured site and protocol characteristics, some part is likely attributable to factors such as the conversion of foreign transactions to US $ using the spot exchange rate at the time of the transaction.

Table II shows two important aspects of trials that impact the costs incurred by an investigator: site work effort and number of patients. Site work effort (SWE) is a patent-pending measure of clinical trial complexity and burden developed by MediData. SWE was constructed as follows. Based on examination of detailed protocols, inclusion and exclusion criteria, the number and use intensity of various procedures such as laboratory tests, blood work, questionnaires and subjective assessments, office consultations and examinations, and use of diagnostic technologies such as x-rays, imaging or heart activity assessments, relative value units (RVUs)[16], or where unavailable or inapplicable, comparable Work Effort Unit (WEUs) created by MediData in conjunction with researchers at the Tufts Center for Study of Drug Development, were assigned to each procedure in a trial protocol. A complexity measure was computed simply as the number of distinct procedures in the trial protocol. An aggregate investigative SWE measure was then computed as the cumulated product of the number and

intensity in use of these procedures, in RVU/WEU units, conducted over the course of the entire protocol for each of the trials.[17] It is important to note that SWE is therefore a protocol-level measure of the work effort required from each site, and that actual resources used by each site in implementing the protocol may differ to some degree. Table II panel (a) reports descriptive statistics for SWE for the 24,236 distinct protocols in our sample. As can be seen from the table , mean and median values of SWE have increased very significantly over time, with the mean value per protocol rising almost three-fold between 1989 and 2011, at AAGRs of 5.2% (mean) and 6.1% (median).[18]

By contrast, as seen in panel (b) of Table II, the mean number of patients per site has fallen substantially over the same period from 25.48 to 11.83, a factor of about two, whereas the median has fallen more dramatically, from 20 to eight. The relative volatility (coefficient of variation) of number patients per site has increased steadily, from about 1.3 in 1989 to about 1.7 in 1999-2000 to 2.0 in 2010, while that for SWE has been relatively stable at about 0.7.

These changes in nominal TGPP, SWE, and number of patients per site suggest that important changes are occurring in the cost and nature of outsourced clinical research. Of course, some of these trends may reflect changing composition of the sample in the mix of therapeutic areas and phases of research, and in the location of sites in the US versus other countries. Tables III through V provide descriptive statistics on the makeup of the sample by various trial characteristics. Table III breaks out the fraction of observations in year by the development phase of the protocol. Clinical trials are conventionally categorized by stage of development. Phase I trials typically enroll a small number of healthy volunteers, and are focused on safety, tolerability, dose-ranging, pharmacokinetics etc. Phase II trials enroll larger numbers of patients and investigate the potential for efficacy by assessing biological activity or effect of the treatment. Phase III trials focus on efficacy of the treatment in therapeutic use, enrolling large numbers of patients. Phase IIIa refers to trials conducted prior to making a submission for regulatory approval, Phase IIIb trial are those initiated after the submission for approval but prior to commercial release. Phase IV trials are conducted after a drug has been approved, often as part of continued investigation of safety. While the fraction of the sample made up by sites involved in Phase I, Phase IIIb, and Phase IV studies was approximately stable, there has been a significant swing in the shares of Phase II and Phase IIIa. In 1989, Phase II trials made up less than 10% of the sample, and Phase IIIa almost 75%. By the end of the

sample period in 2011, Phase II studies comprised almost 30% of trials and Phase IIIa studies dropped to under 60%. If early stage trials are more costly to conduct on a per patient basis, then this shift among trial phases may account for some of the increase in mean TGPP.

Table IV presents the allocation of investigator grants over 15 different therapeutic areas. Reflecting the burden of disease, trials involving the six "largest" therapeutic areas (central nervous system, cardiovascular, respiratory system, endocrine, oncology and anti-infectives) make up 70% of the sample. Shares of central nervous system and oncology trials grew somewhat over time until 2005-6, while cardiovascular shrank, suggesting that to the extent central nervous system and oncology trials are relatively more costly to conduct, these compositional changes may have some effect on increases in average TGPP.

Table V presents the geographic breakdown of the sites in this sample. Over the entire sample time period, 56% of sites were in the US, with most of the remainder in other OECD countries, and only 5.4% in the rest of the world. Interestingly, although the US share is about 80% in both the earliest and latest years, there is substantial year-to-year and trend variability. As shown in Table I we observe considerably smaller numbers of observations in 1989-1991 and 2003-2011 relative to the 1992-2002 time period; any trend analysis is therefore tentative.

III. Hedonic Price Index Methodology

The hedonic pricing approach has a long tradition in economic measurement, going back almost a century.[19] In essence, the hedonic approach treats the item being priced as a bundle of observed characteristics, and using multivariate regression methods, estimates "shadow prices" of each of the observed characteristics and the aggregate price index as a composite of the observed characteristics each multiplied by its shadow price. In practice, given observations in each period *t* on the prices P_{it} of a set of items *i* with characteristics X_{it}, this means estimating a regression model on pooled data of the form $\log(P_{it}) = X_{it}\beta+\gamma Z_t+\varepsilon_{it}$ where Z_t is a set of dummy variables for each period and ε_{it} is a random error term. This semi-log functional form is widely used in hedonic price analysis.[20] Predicted values from this regression provide the basis for computing changes in a "quality-adjusted" composite price index P_t: with a set of time dummies in the regression, the change in the composite index relative to the base period is given by the exponentiated values of their estimated coefficients ($\hat{\gamma}$). Although $E[\exp(P)] \neq \exp(E[P])$ and ε_{it} may not be homoscedastic, suggesting a "smearing" adjustment of the type discussed in the

medical costs literature,[21] with time dummies in the regression these adjustment factors will typically be small. [22] (In the case where residuals are homoscedastic within time periods, adjustments such as the nonparametric method proposed by Duan [1983] will give estimates that are numerically identical to non-adjusted ones. We found very similar adjusted and unadjusted estimated index values, and here we report only estimates with no adjustment for cross-year heteroscedasticity.)

In this application, the "priced item" is the investigator total grant cost per patient. TGPP, and our hedonic regression takes the form $\log(TGPP_{it}) = X_{it}\beta+\gamma Z_t+\varepsilon_{it}$, with X containing site and trial characteristics including planned number of patients at the investigator's site, location and number of sites and countries participating in the trial, phase of development, therapeutic area, and the site work effort (SWE) measure of trial burden and complexity.[23] Z_t are annual indicator variables. Estimated standard errors are Huber-White robust, clustered by trial protocol; computations were carried out in STATA.

IV. Estimation and Price Index Results

We now report results based on various regressions, and calculate corresponding average annual growth rates (AAGRs) of the total grant cost per patient. Although all regressions have as regressors indicator variables for therapeutic class and year, we pool over and then run separate regressions by trial phase; in terms of time periods, we pool over the entire 1989-2011 time period, and then run separate regressions for 1989-1999 and 2000-2011.[24] In all cases the dependent variable is the logarithm of total grant cost per patient (ln TGPP).[25]

Of particular interest to us are the coefficient estimates on two clinical trial characteristics variables—the logarithm of number patients at the site (lpatients) and site work effort (SWE).[26] Note that we have no expectation regarding the sign of the coefficient on the lpatient variable; a negative estimate implies economies of scale, whereas a positive estimate corresponds to diseconomies of scale. Because SWE measures the cumulative burden of various clinical trial protocol procedures, we expect it to have a positive coefficient. Parameter estimates on these two variables, under alternative models and time periods, are presented in Table VI. With two exceptions (both involving lpatients in Phase II trials), all of the estimated coefficients are statistically significant at the 1% level, based on robust standard errors.

A number of results are striking. As shown in the top panel of Table VI, when pooled over all phases (but including trial phase as indicator variables), globally and for ex-US (Rest of World), in all three time period regressions the estimated coefficient on lpatients is positive and highly significant; however, for the US Only model the coefficient estimate is negative and significant. The implied estimated elasticities of TGPP with respect to patients range from -0.122 to 0.183.

The pattern of estimates on lpatients becomes a bit more nuanced when separate regressions are run by trial phase. Specifically, a general pattern that prevails is that negative estimates occur for the Phase I and Phase II regressions for the All and Rest of World regressions, but these estimates become positive and ever larger as one moves to the increasingly larger patient size. In Phase IIIA, Phase IIIB and Phase IV trials, almost all the estimates are positive and significant even at p-values <0.01. Also notable is the substantial range in estimates of the elasticity of TGPP with respect to patients, from -0.176 to 0.320 in the pooled 1989-2011 regressions, even larger from-0.219 to 0.305 in the 1989-1999 regressions, and in the 2000-2011 regressions, ranging from -0.190 to 0.272.

A second set of striking findings in Table VI is that every one of the estimates on the SWE variable is positive and statistically significant at p-values < 0.01, with the general (but not quite universal) pattern being that the positive estimates increase monotonically as one moves from the small Phase I to the larger Phase IIIB and Phase IV trials. The steepness of the positive slope with larger trial phase is flatter for the US Only regressions, however, than for the All and Rest of World regressions, with the Phase IV All and Rest of World estimates being particularly large; the vast majority of estimates on the SWE variable are in the range of 0.01 to 0.03. Using a mean value of SWE of about 25 (see Table II), a one-unit increase in SWE changes it by about 4% (1/25), leading to about on average a 2% increase in TGPP, suggesting an elasticity of TGPP with respect to SWE at about 0.50 (= 0.02/0.04). That is a very substantial effect.

A third implication of findings in Table VI is that they help explain factors affecting increases in trial costs. Specifically, as SWE has increased over time and number of patients has decreased, particularly in the US where for each phase coefficients on lpatients are mostly negative, TGPP is increased. Whether the changing composition among trial phases (towards Phase II and away from Phase IIIA—see Table III) can “explain” the increase in TGPP merits further examination.

We now move on to consider implications of these various regression models for the growth rate of our price indexes. As discussed above, annual values of an hedonic price index can be constructed from estimated coefficients on indicator variables for year. We summarize the growth rate of this index by computing the Annual Average Growth Rate (AAGR), which is the mean of year-on-year percentage changes in the index values.[27] In the top panel of Table VII we report estimates of AAGRs in a "base" model that excludes our two prominent quality measures, namely, lpatients and SWE, which from Table VI we have observed as being highly statistically significant. To quantify the importance of including these trial site characteristics in our hedonic regression equation, we compare AAGRs of predicted TGPP with and without the lpatient and SWE variables included by examining the relative growth of coefficients on the yearly indicator variables. The results are quite striking. With the pooled 1989-2011 regression, relative to the base model, TGPP grows much more slowly when the trial site characteristics are included—4.31%/6.96% for All (38% lower AAGR), 3.62%/7.01% for US Only (48% smaller AAGR), and 6.05%/8.70% for Rest of World (30% lower AAGR). For the 1989-1999 regressions (second last column), the corresponding percent reductions in AAGRs are more modest—6% All, 31% US Only, and 19% for Rest of World, but for the most recent 2000-2011 time period regressions (last column), they are not only large proportionately—40% lower AAGR for All, 56% for US Only, and 28% for Rest of World, but the absolute differences in AAGRs are substantial—3.95% (9.93% - 5.98%) for All, 3.52% (12.36% - 8.84%) for Rest of the World, and 4.10% (7.38% - 3.28%) for the US Only. We conclude, therefore, that controlling for the clinical trial quality characteristics lpatient and SWE results in much lower AAGRs, and helps explain in part why it is that TGPP has been increasing steadily over the last two decades.

The bottom three panels of Table VII report AAGRs based on separate regressions by trial phase. As seen in the first column of Table VII based on pooled 1989-2011 regressions, over all trial phases US Only AAGRs are smaller than the All regression AAGRs, with the Rest of World regression AAGRs being greater than those in All for Phases I, II, Phases IIIA, and IIIB, but less than All for Phase IV. The variation in AAGRs within each set of regressions is quite large—from 3.78% to 11.91% in All, 3.48% to 6.78% in US Only, and 6.29% to 15.51% in Rest of World regressions. AAGRs are generally lowest in Phase IIIA and IIIB, and mostly highest in Phase II and Phase IV. Even though the regressions involve pooled 1989-2011 data,

as seen in the second and third column there is considerable variation across the two time intervals within the pooled regression.

Comparing 1989-1999 AAGRs from the pooled regression (column two) with those from the separate 1989-1999 regression (column four), and the 1999-2011 AAGRs from the pooled regression (column three) with those from the separate 2000-2011 regression (last column) provides some evidence regarding parameter stability. The 1989-1999 relative rankings of AAGRs across trial phases in columns two and four is quite robust, but slightly less so when comparing relative rankings in columns three and five. Particularly notable is the uniformly greatest growth rate during the 2000s in Phase IV trials, with substantial but less uniformly large AAGRs in Phase I studies.

Tables VIII and IX present results of two sets of exploratory findings. Although issues of sample size are likely to become important, we estimate ln TGPP equations at the level of the therapeutic class, pooled 1989-2011 and separately for 1989-1999 and 2000-2011; as before we do three geography-based estimations—All, US-Only and Rest of World. There are 15 therapeutic classes in our trial data, 14 of them involving biopharmaceuticals plus a devices and diagnostics category. In Table VIII we report AAGRs by therapeutic class for the US-Only and Rest of World regressions; note that because of the absence of any observations in some years, there is some variability from the 1989-2011, 1989-99 and 2000-11 beginning and ending years, as is described at the bottom of the table. The most striking feature of Table VIII is the substantial variability in the AAGRs; not shown is the even greater variability in the estimates on the year indicator variables within each therapeutic class. While is it likely that there is in fact substantial variation across therapeutic classes in the rate of change of trial costs, some of the variation is likely attributable to to the smaller sample sizes in certain years that result from disaggregating into 15 therapeutic classes. This makes it difficult to estimate the hedonic index values precisely, and particularly outside the US there are only enough observations for some therapeutic classes to estimate the index values in a limited number of years.

We conclude that constructing price indexes for clinical trials at the level of therapeutic classes (in our case, which number 15) is likely to be infeasible because of sample size issues, particularly for ex-US sites.

Our final exploratory price index analysis involves aggregating up from individual multiple sites within a given trial to the trial level at which there is a common protocol. This

allows us to examine whether number of trial sites and the geographical scope of the sites affects our dependent variable, ln TGPP. This aggregation reduces our sample size from the 207,950 sites in Table VI to 24,172 distinct trials. We construct two new variables that vary at the level of the individual trial protocol: number of sites, and number of sites per country. We also recalculate the dependent variable, lpatient and the SWE variable at the trial level of aggregation. We do not have a prior expectation regarding the sign of the coefficient on total number of sites per trial. This coefficient will capture whether or not there are cost impacts (at the site level) of allocating a given number of patients across different numbers of sites. To the extent there are fixed costs incurred at each site for setting up patient recruitment, independent of the number of patients enrolled at a given site, then holding the numbers of patients constant the aggregate TGPP would be expected to increase with the number of sites. On the other hand, if fixed costs are largely trial specific rather than site-specific, and are carried in the "overhead" part of trial costs which we do not observe in these data, then they will either not affect site-level costs, i.e. no observable impact of number of sites on aggregate TGPP, or to the extent that they reduce site-specific costs otherwise borne by investigators, will result in a negative relationship between aggregate TGPP and number of sites.

Some of these trial-level fixed costs are likely to be country-specific, reflecting factors such as national institutional review boards, import duties and tariffs, medical licensing conventions or other costs of conforming to a given country's regulatory framework and infrastructure. To the extent that these costs are "pushed down" to individual sites, rather than absorbed in the overall "overhead" cost of the trial then aggregate TGPP may be affected by the number of sites per country. We therefore also control for each trial's number of sites per country.

In Table IX, we report coefficient estimates on the number of sites, and the number of sites per country, for regressions at two levels of aggregation: Pooled over phases (but with phase indicator variables included as regressors), and separately by trial phase.[28] When pooled over phases, the estimate on number of sites is positive and strongly significant, while the number of sites per country is negative but not significant. When estimated separately by phase, signs on the number of sites variable are mixed but monotonically decline moving from early to late phases, and although none is statistically significant. However, all but one of the estimates on the number of sites per country are positive, statistically significant in the case of Phase II and

IIIA trials. In almost all cases, however, the absolute magnitudes of the coefficient estimates are very small—an order of magnitude or smaller than those on the lpatient and SWE variables reported in Table VI. We conclude that at the level of a clinical trial protocol, the number of sites and number of countries per site do not appear to have a material effect on the total grant cost per patient. These trial characteristics might, however, have varying effects on the sponsors' overall "headquarters" overhead costs, which we do not observe.

V. Summary, Conclusions, Limitations and Future Research

Expenditures on clinical trials required to develop new drugs have increased dramatically over the past 30 years. To better understand the underlying causes it is important to be able to decompose increases in total spending into the "price effect", the "quantity effect," and the "quality" effect. Are biopharmaceutical companies doing more clinical research, or has the cost of doing a given amount of research increased, or are both occurring? In this study we focus on the "unit costs" of some aspects of conducting clinical trials. These have risen substantially in recent decades, outpacing general inflation and other measures of changes in costs of other inputs to biomedical R&D. Our results suggest that these increases in trial costs are not solely attributable to changes in input costs such as wages, equipment, and facilities. They also appear to have been driven to a substantial extent by two other phenomena: smaller numbers of patients per site, and increases in the "effort" level required by investigators as study protocols have required more costly and complex monitoring and testing of subjects. While these in turn are driven to some extent by cost differences across therapeutic classes and phases of clinical development, the effects we find are estimated controlling for such study characteristics, and are not just an artifact of changes in the composition of the sample. The size of the effects that we find implies that any effort to track costs of clinical research should pay close attention to the nature of study protocols and the organization and management of trials. These findings point to the value of using the hedonic regression methodology in this context.

The price indexes for commercial clinical research constructed here appear to behave very differently from those computed by NIH for input costs for public sector biomedical R&D, and may merit more careful attention by government statistical agencies and other entities with an interest in tracking R&D costs in this sector. The AAGR of a price index that controls for therapeutic class and phase of development grows almost twice as fast as the NIH BRDPI input costs index. Interestingly, once the scale of investigator/site activity and the effort required by a

study protocol are also controlled for, the estimated "quality-adjusted" rate of inflation within the US is quite similar to the BRDPI. This suggests that increases in commercial clinical trial costs are driven primarily by changes in the nature of clinical research rather than by inflation in input costs. Commercial databases such as the one we have used here appear to have great potential as a source of data for such price index measurement purposes. Using these data it would appear to be feasible to reliably compute measures of price inflation for this aspect of clinical research, and to do this separately for different phases of clinical development, and for some but not all therapeutic classes. The geographic reach of these data sources also presents interesting opportunities to benchmark R&D costs across different regions and countries.

There are some important limitations to our study. In particular, we look only at one component of trial costs: payments to clinical investigators. In this dataset, these account for about one half of the total cost of a trial. It may well be that some of the higher per-patient costs created by having fewer patients per site and increased effort required by the protocol are offset by savings in the costs of centralized administration and co-ordination of trials that we do not observe here. Limited availability of data prevents us from drawing strong conclusions about trends in total trial costs and underlying factors in recent years. Further, it is unclear how well the measure of "site work effort" used here captures differences in the burden imposed by, for example, running more complex trial protocols, as opposed to use of more costly interventions or methods of measuring endpoints. Lastly, since the identity of study sponsors and investigators was not available to us, we were not able to investigate differences in costs across (for example) trials sponsored by large versus small commercial entities, or where public sector or non-profit organizations are involved as sponsors or investigators.

We look forward to addressing these questions in future work.

References

Azoulay, Pierre [2004], "Capturing Knowledge Within and Across Firm Boundaries: Evidence from Clinical Development", *American Economic Review* 94(5):1591-1612, December.

Azoulay, Pierre and Ariel Fishman [2008], "Doctors, $$ and Drug Development: The Rise for For-Profit Experimental Medicine", paper presented at the National Bureau of Economic Research Conference, The Location of Biopharmaceutical Activity, Savannah, Georgia, March 7-8.

Berndt, Ernst R. [1991], *The Practice of Econometrics: Classic and Coontemporary*, Reading, MA: Addison-Wesley Publishing Company.

Berndt, Ernst R., Zvi Griliches and Neal J. Rappaport [1995], "Econometric Estimates of Price Indexes for Personal Computers in the 1990s", *Journal of Econometrics* 68(1):243-68, July.

Berndt, Ernst R. and Neal J. Rappaport [2001], "Price and Quality of Desktop and Mobile Personal Computers: A Quarter Century Historical Overview", *American Economic Review* 91(2):268-73, May.

Corrado, Carol [2008]. *Report on Workshop on Developing a New Research Data Infrastructure for the Study of Organizations and Innovations*. Arlington, Virginia: National Science Foundation Headquarters, July 23, 2008. New York: The Conference Board, November 10.

Cutler, David M., Mark McClellan, Joseph P. Newhouse and Dahlia K. Remler [2001], "Pricing Heart Attack Treatments", ch. 8 in David M. Cutler and Ernst R. Berndt, eds., *Medical Care Output and Productivity*, Chicago IL: University of Chicago Press for the National Bureau of Economic Research, Studies in Income and Wealth Vol. 62, 305-47.

Cutler, David M., Mark McClellan, Joseph P. Newhouse and Dahlia K. Remler [1998], "Are Medical Prices Declining? Evidence from Heart Attack Treatments", *Quarterly Journal of Economics* 113:991-1024, November.

DiMasi, Joseph A., Ronald W. Hansen and Henry G. Grabowski [2003], "The Price of Innovation: New Estimates of Drug Development Costs", *Journal of Health Economics* 22(2):151-85, March.

Duan, Nathan [2001], "Smearing Estimate: A Nonparametric Retransformation Method", *Journal of the American Statistical Association* 78: 605-10.

Fast Track Systems [2006]. Grants Manager Data Dictionary. Conshohocken, PA: Fast Track Systems, Inc. Excel spreadsheet file.

Getz, Kenneth A., Julia Wenger, Rafael A. Campo, Edward S. Seguine and Kenneth L. Kaitin [2008], "Assessing the Impact of Protocol Design Changes on Clinical Trial Performance", *American Journal of Therapeutics* 15(5):450-57.

Manning, Willard G. [1998], "The Logged Dependent Variable, Heteroskedasticity, and the Retransformation Problem", *Journal of Health Economics* 17:283-95.

Manning, Willard G. and John Mullahy [2001], "Estimating log models: To Transform or Not to Transform?", *Journal of Health Economics* 20:461-94.

MediData Solutions Worldwide, Inc. [2007]. "MediData Solutions Partners with Fast Track Systems to Accelerate Enterprise-Wide Deployment of Electronic Clinical Trials." Press release, February 7. http://www.mdsol.com/documents/press/20070207.pdf, accessed December 7, 2008.

Mullahy, John [1998], "Much Ado About Two: Reconsideration Retransformation and the Two-Part Model in Health Econometrics", *Journal of Health Economics* 17:247-81.

National Institutes of Health, Office of Budget [2011], "Biomedical Research and Development Price Index (BRDPI): Fiscal Year 2010 Update and Projections for FY 2011-FY 2016", January 24. Available online at http://www.officeofbudget.od.nih.gov/pdfs.../BRDPI_Proj_2011_Final.pdf.

Pharmaceutical Research and Manufacturers of America [2011], *Pharmaceutical Industry Profile 2011*, Washington DC: PhRMA, April.

Schultze, Charles L. and Christopher Mackie, eds. [2002], *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes*, Washington DC: National Academy Press for the National Academy of Sciences.

Triplett, Jack [2006] *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*, Paris: OECD.

Table I - Grant Total per Patient

year	mean	p50	sd	N
1989	3772.59	2779.43	2921.34	1370
1990	4385.77	3147.77	4016.57	3443
1991	3774.43	2774.83	4186.06	9288
1992	3493.63	2399.00	6129.31	14126
1993	3664.58	2325.45	4254.28	15733
1994	3911.39	2882.35	3925.46	16625
1995	4183.00	3203.85	3941.90	15670
1996	4884.89	3748.77	4708.14	14442
1997	4549.12	3200.00	4422.39	13321
1998	5393.70	3948.38	5445.89	14370
1999	5501.08	4361.94	4874.07	13943
2000	6220.42	4682.79	6243.02	18671
2001	6078.96	4777.00	5150.11	16864
2002	6567.58	4744.10	5984.32	12201
2003	8147.90	6765.00	6866.55	6515
2004	10264.00	8582.72	7758.22	3216
2005	11412.77	9682.02	7828.89	2693
2006	12364.68	10900.00	7460.17	4012
2007	13001.19	10738.47	8863.90	4764
2008	14834.64	12720.94	10328.42	3216
2009	16518.28	13965.42	12550.80	4591
2010	15099.19	12581.93	10860.27	4814
2011	16566.55	13222.14	13556.92	2188
Total	6191.80	4195.07	6860.91	216076

Table II

	(a) Site Work Effort - SWE (Protocol level)				(b) Patients (Site level)			
YEAR	MEAN	MEDIAN	STD. DEV.	N	MEAN	MEDIAN	STD. DEV.	N
1989	17.07	13.29	12.88	240	25.48	20.00	33.2476	1370
1990	16.94	13.44	13.38	512	21.26	15.00	27.3178	3443
1991	14.85	11.16	12.11	1280	20.64	15.00	32.4249	9288
1992	17.74	13.46	15.45	1833	18.50	12.00	25.253	14126
1993	18.92	13.66	17.79	2024	17.04	12.00	24.5758	15733
1994	21.58	16.22	18.20	1967	16.37	12.00	24.6318	16625
1995	23.41	18.11	21.20	1826	14.39	10.00	26.52	15675
1996	23.69	18.64	19.32	1791	14.83	10.00	36.2925	14444
1997	24.93	19.26	20.11	1540	14.14	10.00	17.6792	13321
1998	25.77	19.64	19.81	1495	14.60	10.00	24.9971	14370
1999	26.87	20.55	23.89	1369	13.85	10.00	17.0209	13944
2000	28.18	22.44	22.47	1890	11.27	9.00	14.7171	18696
2001	29.91	22.78	24.18	1920	12.11	10.00	13.8476	16907
2002	28.91	21.97	24.06	1389	12.07	10.00	12.2206	12250
2003	35.46	29.29	27.26	933	13.30	10.00	18.5263	6629
2004	43.57	34.33	34.09	569	12.18	9.00	14.5193	3362
2005	47.65	39.06	38.04	307	10.73	10.00	14.041	2784
2006	45.53	35.62	35.96	274	11.80	10.00	9.62717	4027
2007	37.71	30.58	22.10	227	10.30	9.00	8.47037	4784
2008	48.64	41.55	37.49	183	11.10	7.00	19.9492	3291
2009	48.98	41.76	33.86	263	11.81	8.00	15.4748	4729
2010	46.16	38.52	33.39	275	11.79	8.00	23.6707	4907
2011	45.69	41.31	30.59	129	11.83	8.00	17.0184	2224
Total	25.94	19.21	23.28	24236	14.42	10.00	22.58	216929

Table III – Development Phase Percentages

Year	Phase I	Phase II	Phase IIIA	Phase IIIB	Phase IV
1989	2.12%	9.12%	72.55%	6.35%	9.85%
1990	3.08%	13.94%	69.53%	5.84%	7.61%
1991	3.45%	15.13%	50.67%	12.75%	18.01%
1992	3.58%	14.14%	58.47%	6.98%	16.82%
1993	3.96%	18.34%	56.79%	4.74%	16.16%
1994	3.70%	18.80%	58.69%	7.65%	11.16%
1995	4.20%	20.52%	54.23%	6.63%	14.41%
1996	4.84%	23.09%	52.94%	8.55%	10.58%
1997	4.14%	18.08%	56.53%	8.85%	12.40%
1998	3.63%	19.08%	49.45%	13.84%	13.99%
1999	4.06%	20.35%	49.25%	14.08%	12.26%
2000	3.39%	17.59%	49.86%	17.72%	11.44%
2001	3.64%	16.60%	50.21%	14.95%	14.60%
2002	3.00%	20.99%	46.78%	16.56%	12.68%
2003	4.06%	17.79%	45.26%	16.52%	16.38%
2004	4.49%	23.38%	42.15%	18.59%	11.39%
2005	2.33%	26.08%	51.33%	9.05%	11.21%
2006	3.25%	15.50%	69.23%	3.35%	8.67%
2007	2.17%	25.90%	49.85%	15.49%	6.58%
2008	2.98%	23.06%	48.40%	16.38%	9.18%
2009	3.45%	23.41%	58.64%	7.76%	6.75%
2010	2.45%	31.97%	46.06%	11.98%	7.54%
2011	1.66%	25.99%	55.35%	6.74%	10.25%
Total	3.67%	19.26%	53.07%	11.18%	12.83%

Table entries are the fraction of investigator contracts in that year for studies at each phase of clinical development. Based on 216,929 total observations.

Table IV - Share Distribution by Therapeutic Class

Year	Anti-Infective	Cardio-vascular	Central Nervous System	Dermatology	Devices and Diagnostics	Endocrine	Gastro-intestinal
1989	7.66%	14.82%	14.38%	4.09%	0.00%	3.28%	10.07%
1990	9.09%	21.64%	9.70%	2.27%	0.20%	7.29%	9.00%
1991	8.81%	18.62%	14.69%	2.45%	0.12%	6.01%	13.18%
1992	8.10%	16.69%	15.33%	2.63%	0.27%	5.88%	9.80%
1993	10.13%	19.70%	14.37%	2.58%	0.35%	7.29%	9.92%
1994	11.56%	20.82%	16.72%	2.15%	0.48%	6.42%	6.35%
1995	8.09%	18.65%	16.66%	2.07%	0.05%	6.46%	4.43%
1996	8.42%	16.45%	17.50%	1.88%	0.18%	7.57%	3.82%
1997	6.58%	15.54%	19.06%	1.73%	0.29%	9.10%	4.51%
1998	7.24%	15.85%	15.49%	2.75%	0.15%	11.69%	1.88%
1999	5.49%	18.89%	14.08%	2.36%	0.32%	11.93%	2.76%
2000	6.93%	15.28%	14.85%	2.40%	0.52%	13.61%	3.36%
2001	10.56%	9.39%	13.95%	2.83%	0.15%	14.83%	5.20%
2002	6.39%	8.42%	13.84%	2.02%	0.17%	17.37%	3.85%
2003	8.63%	9.94%	22.85%	6.14%	0.86%	12.25%	2.16%
2004	5.68%	12.31%	27.96%	1.81%	1.04%	13.06%	0.39%
2005	4.92%	21.62%	21.12%	3.84%	1.11%	5.28%	2.12%
2006	2.36%	12.14%	27.04%	2.48%	0.15%	19.44%	3.30%
2007	4.77%	4.52%	17.52%	1.17%	0.31%	18.42%	7.46%
2008	3.77%	6.05%	17.23%	3.80%	0.64%	21.27%	5.62%
2009	9.45%	1.69%	18.95%	1.78%	2.24%	23.77%	4.95%
2010	6.64%	1.26%	12.43%	5.05%	2.71%	31.57%	4.18%
2011	1.35%	0.63%	14.88%	6.07%	6.03%	28.46%	0.05%
Total	7.87%	14.79%	16.21%	2.56%	0.47%	11.44%	5.29%

Table entries are the fraction of investigator contracts in that year for studies in each therapeutic area. Based on 216,929 total observations

Table IV (con't.) - Share Distribution by Therapeutic Class

Year	Genitourinary System	Hematology	Immuno-modulation	Oncology	Ophthal-mology	Pain and Anesthesia	Pharmaco-kinetics	Respiratory System
1989	10.51%	0.58%	18.76%	5.84%	1.68%	1.61%	1.68%	5.04%
1990	5.14%	1.95%	7.75%	8.51%	8.19%	0.46%	1.54%	7.26%
1991	5.33%	0.39%	8.57%	4.94%	2.57%	1.07%	2.02%	11.24%
1992	8.58%	0.42%	6.26%	5.12%	0.68%	3.27%	2.10%	14.87%
1993	9.76%	0.70%	6.06%	3.37%	0.66%	0.90%	2.07%	12.15%
1994	5.73%	0.94%	6.09%	4.84%	1.41%	0.87%	2.36%	13.27%
1995	7.45%	0.54%	5.33%	8.22%	1.07%	1.33%	2.29%	17.34%
1996	4.94%	1.70%	7.55%	9.62%	0.80%	2.85%	2.51%	14.21%
1997	6.51%	0.62%	5.33%	10.81%	1.70%	1.37%	2.67%	14.17%
1998	8.20%	1.75%	5.82%	10.51%	1.34%	1.51%	2.39%	13.43%
1999	6.66%	3.13%	10.23%	10.44%	1.78%	2.20%	2.57%	7.16%
2000	5.60%	3.29%	7.05%	11.26%	0.59%	2.12%	1.87%	11.27%
2001	7.58%	1.41%	9.30%	10.46%	1.20%	2.98%	1.87%	8.28%
2002	9.82%	1.89%	9.63%	10.21%	1.52%	4.58%	1.56%	8.73%
2003	4.89%	2.07%	6.46%	10.47%	1.28%	4.16%	1.89%	5.96%
2004	8.36%	3.84%	9.52%	10.14%	0.54%	0.27%	2.23%	2.86%
2005	2.26%	6.07%	7.79%	18.25%	0.36%	1.51%	1.01%	2.73%
2006	3.43%	0.30%	8.12%	11.12%	1.04%	6.48%	0.60%	1.99%
2007	11.64%	0.65%	9.49%	8.38%	0.92%	7.34%	0.10%	7.32%
2008	5.89%	0.58%	15.16%	14.80%	2.25%	2.22%	0.21%	0.52%
2009	7.74%	1.61%	7.70%	14.55%	1.46%	1.97%	0.68%	1.48%
2010	5.34%	1.87%	5.24%	11.98%	2.65%	6.07%	0.79%	2.20%
2011	3.06%	10.12%	5.35%	12.55%	6.07%	1.17%	2.52%	1.71%
Total	6.99%	1.62%	7.43%	9.00%	1.40%	2.35%	1.98%	10.59%

Table V -Sites Percentage

	(a) US vs. Ex-US		(b) Ex-US	
Year	US	RoW	OECD	Other
1989	81.82%	18.18%	100.00%	0.00%
1990	77.17%	22.83%	99.11%	0.89%
1991	67.51%	32.49%	99.83%	0.17%
1992	49.19%	50.81%	99.79%	0.21%
1993	47.31%	52.69%	99.59%	0.41%
1994	46.57%	53.43%	99.64%	0.36%
1995	47.50%	52.50%	99.54%	0.46%
1996	55.23%	44.77%	99.72%	0.28%
1997	48.18%	51.82%	98.87%	1.13%
1998	46.09%	53.91%	98.06%	1.94%
1999	53.49%	46.51%	96.62%	3.38%
2000	56.33%	43.67%	92.05%	7.95%
2001	56.54%	43.46%	89.06%	10.94%
2002	54.08%	45.92%	90.19%	9.81%
2003	52.69%	47.31%	90.66%	9.34%
2004	54.46%	45.54%	84.91%	15.09%
2005	62.75%	37.25%	83.32%	16.68%
2006	85.10%	14.90%	83.50%	16.50%
2007	87.54%	12.46%	58.89%	41.11%
2008	68.82%	31.18%	55.36%	44.64%
2009	77.65%	22.35%	62.25%	37.75%
2010	76.56%	23.44%	68.87%	31.13%
2011	68.03%	31.97%	53.59%	46.41%
Total	55.65%	44.35%	94.61%	5.39%

Table entries are the fraction of investigator contracts in that year located in each geographic area. Panel (a) shows the breakdown between US and all other countries. Panel B breaks out the Ex-US countries into OECD member countries (Canada, Austria, Belgium, Ireland, Luxembourg, Monaco, Netherlands, Switzerland, Finland, France, Norway, Germany, Italy, Spain, Sweden, Denmark, Japan, Australia, New Zealand, United Kingdom) versus all others. Based on 216,929 total observations.

TABLE VI: PARAMETER ESTIMATES ON LOG PATIENTS AND SITE WORK EFFORT TRIAL CHARACTERISTICS VARIABLES

	Pooled 1989-2011		1989-1999		2000-2011	
	lpatients	SWE	lpatients	SWE	lpatients	SWE
Phases Pooled						
All	0.148	0.0205	0.183	0.0279	0.0419	0.0152
US Only	-0.122	0.0171	-0.120	0.0219	-0.117	0.0143
Rest of World	0.124	0.0225	0.128	0.0296	0.0541	0.0161
By Phase						
All						
Phase I	-0.145	0.0130	-0.168	0.0164	-0.0966	0.00972
Phase II	-0.0173	0.0182	-0.011†	0.0222	-0.0321	0.0149
Phase IIIA	0.120	0.0198	0.173	0.0285	-0.0240	0.0137
Phase IIIB	0.155	0.0284	0.226	0.0341	0.0299	0.0244
Phase IV	0.320	0.0344	0.305	0.0489	0.272	0.0253
By Phase						
US Only						
Phase I	-0.176	0.0143	-0.219	0.0194	-0.104	0.0102
Phase II	-0.164	0.0159	-0.197	0.0173	-0.125	0.0144
Phase IIIA	-0.102	0.0164	-0.0986	0.0239	-0.0940	0.0127
Phase IIIB	-0.0631	0.0257	-0.0541	0.0238	-0.0513	0.0263
Phase IV	-0.170	0.0247	-0.153	0.0273	-0.190	0.0217
By Phase						
Rest of World						
Phase I	-0.120	0.0109	-0.131	0.0134	-0.0799	0.00788
Phase II	-0.01	0.0186	-0.0313	0.0265	-0.0080†	0.0143
Phase IIIA	0.0826	0.0220	0.121	0.0277	-0.0588	0.0157
Phase IIIB	0.101	0.0276	0.115	0.0372	0.0409	0.0213
Phase IV	0.217	0.0482	0.171	0.0619	0.225	0.0322

Notes: Table entries are the estimated coefficients on lpatients and SWE in a regression of log(TGPP) on these and other explanatory variables. With the exception of the coefficients marked †, all coefficients in the table were statistically distinguishable from zero at the $p<0.01$ level using robust standard errors, clustered by trial protocol. The phases pooled regression includes a constant, indicator variables for therapeutic class, trial phase, and years. The by phase regressions include a constant and indicator variables for therapeutic class and year. Number of observations was 207,950 in the top panel All regressions, 118,477 in the US Only regressions, and 89,473 in the Rest of World regressions.

TABLE VII: AVERAGE ANNUAL GROWTH RATES OF CLINICAL TRIAL COSTS: ALTERNATIVE MODELS AND TIME PERIODS

	1989-2011 Regression AAGR			1989-1999	2000-2011
	1989-2011	1989-1999	1999-2011	Regression AAGR	Regression AAGR
Base Model^					
All	6.96%	3.98%	9.45%	3.79%	9.93%
US Only	7.01	5.80	8.01	5.70	7.38
Rest of World	8.70	6.66	10.39	6.54	12.36
Add SWE and lpatients^					
All	4.31	3.96	4.59	3.54	5.98
US Only	3.62	4.28	3.07	3.92	3.28
Rest of World	6.05	6.02	6.08	5.29	8.84
With SWE and lpatients by phase^^					
All					
Phase I	7.48%	5.67%	8.98%	5.19%	9.08%
Phase II	6.01	6.58	5.54	6.91	5.98
Phase IIIA	4.04	3.88	4.18	3.25	5.16
Phase IIIB	3.78	2.32	5.00	1.79	5.65
Phase IV	11.91	9.25	14.13	8.17	15.75
US Only					
Phase I	6.78	5.72	7.66	5.07	8.58
Phase II	6.00	7.77	4.52	7.61	4.32
Phase IIIA	3.78	4.33	3.32	3.68	2.93
Phase IIIB	3.48	3.92	3.11	4.24	3.34
Phase IV	6.43	8.02	5.11	7.80	4.03
Rest of World					
Phase I	15.51	11.69	18.70	10.62	23.96
Phase II	7.64	8.48	6.95	8.65	9.52
Phase IIIA	6.29	6.87	5.81	6.30	7.87
Phase IIIB	12.29	15.54	9.59	13.72	11.55
Phase IV	9.36	10.73	8.22	8.87	11.74

Notes: Table entries are the annual average growth rate (AAGR) of an hedonic price index constructed from the estimated coefficients on indicator variables for year. See text. ^Regressions also include constant and indicator variables for therapeutic class, and trial phase. ^^Regressions also include constant and indicator variables for therapeutic class.

TABLE VIII: AVERAGE ANNUAL GROWTH RATES OF CLINICAL TRIAL COSTS BY THERAPEUTIC AREA AND TIME PERIOD

Therapeutic Class	1989-2011 Regression 1989-2011 AAGR	1989-99 Regression 1989-99 AAGR	2000-2011 Regression 2000-2011 AAGR
		US Only Regressions	
Anti-infective	4.73%	0.81%	7.65%
Cardiovascular	18.64	4.95	14.10
Central Nervous System	5.05	5.00	5.35
Dermatology	6.56	9.18	2.10
Devices & Diagnostics	17.62[a]	18.32[b]	22.08
Endocrine	4.94	5.60	4.20
Gastrointestinal	7.11	5.77	3.14
Genitourinary System	9.02	9.96	8.87
Hematology	11.17	25.24	2.71
Immunomodulation	6.08	6.77	5.82
Oncology	6.70	3.81	6.51
Ophthalmology	9.67	8.96	21.97
Pain & Anaesthesia	10.07	9.59	11.25
Pharmacokinetics	6.84	5.27	8.98
Respiratory System	8.56	5.84	12.13

Therapeutic Class	1989-2011 Regression 1989-2011 AAGR	1989-99 Regression 1989-99 AAGR	2000-2011 Regression 2000-2011 AAGR
		Rest of World Regressions	
Anti-infective	10.19%	12.66%	14.13%
Cardiovascular	15.88	13.24	13.95
Central Nervous System	9.68	6.18	13.32
Dermatology	15.99[c]	15.06	-1.41
Devices & Diagnostics	17.50[d]	-9.97 [g]	30.61[h]
Endocrine	6.10	7.66	5.83
Gastrointestinal	20.14	3.27	23.55
Genitourinary System	6.58	4.75	9.01
Hematology	11.40[e]	19.24	2.38[i]
Immunomodulation	13.17	18.07	12.01
Oncology	9.60	6.39	14.40
Ophthalmology	29.60	42.85	17.69
Pain & Anaesthesia	43.98[f]	47.61	45.37
Pharmacokinetics	7.99[f]	1.65	12.93[k]
Respiratory System	11.52	10.25	20.09[k]

Notes: [a]1999-2011; [b]1990-94; [c]1990-2003; [d]1992-98; [e]1989-2005; [f]1989-2006;[g]1993-99; [h]2009-11; [i]2000-5; [k] 2000-06

TABLE IX: PARAMETER ESTIMATES IN MODEL ESTIMATED AT TRIAL LEVEL WITH ADDITIONAL TRIAL-SPECIFIC EXPLANATORY VARIABLES, POOLED 1989-2011 DEPENDENT VARIABLE: ln TGPP

	No. Sites	No. Sites/Country	Sample Size
Pooled Over Phases	0.00164***	-0.000979	24,172
By Phase			
Phase I	0.0223	0.0102	5,557
Phase II	0.00116	0.00506***	5,775
Phase IIIA	0.000371	0.00161*	8,953
Phase IIIB	-0.000837	0.000454	1,735
Phase IV	-0.00206	-0.00126	2,152

Notes: ***, ** and * denote statistical significance at p-values of 0.01, 0.05 and 0.10, respectively. The phases pooled regression includes a constant, lpatient, SWE, and indicator variables for therapeutic class, trial phase and years. The by phase regressions include a constant, lpatient, SWE, and indicator variables for therapeutic class and year. Regressions are pooled into All, 1989-2011.

ENDNOTES

[1]Pharmaceutical Research and Manufacturers of America [2011], Appendix Table 5, p. 45. Expenditure on Phase 1 trials was $3.753 billion, Phase 2 $7.124 billion, Phase 3 $16.300 billion, and Phase 4 5.303 billion.
[2] See http://report.nih.gov/rcdc/categories/default.aspx, visited 11/30/2011.
[3] Total R&D spending by PhRMA member companies was $26.0 billion in 2000, and $49.4 billion in 2010, an increase of 90%. Pharmaceutical Research and Manufacturers of America [2011], Appendix Table 1, p. 42.
[4] The total NIH budget obligations in fiscal year 2000 was $l7.8 billion, and $31.0 billion in fiscal year 2010, an increase of 74%. National Institutes of Health, Mechanism Detail, Actual Obligations, available at http://officeofbudget.od.nih.
[5] Getz, Wenger, Campo et al. [2008].
[6] Cutler, McClellan, Newhouse and Remler [1998,2001].
[7] Corrado [2008].
[8] See, for example, Berndt and Rappaport [2001] and the references cited therein.
[9] For more detailed discussion, see ch. 4 in Schultze and Mackie [2002].
[10] Azoulay [2004], and Azoulay and Fishman [2008].
[11] National Institutes of Health [2011], p. 1. In index number nomenclature, the BRDPI is annual chained Laspeyres price index.
[12] The dataset was originally compiled by Fast Track Systems; MediData acquired Fast Track in 2007. See Fast Track Systems [2006] and MediData Solutions Worldwide [2007] for further details.
[13] These AAGRs are literally the arithmetic means of year-over-year growth rates; the compounded average annual growth rate is slightly less, at 6.7%; see note 27 below.
[14] In the somewhat larger sample (245,803 records) that includes contracts where the number of patients is not specified, figures are very similar: mean TGPP rises from $3,752 to $15,567 at an AAGR of 7.1%.
[15] National Institutes of Health [2011], Supplemental Table A, p. 6.
[16] RVUs are measures constructed by Medicare to estimate the relative level of physician time, skill, training, and expertise and required equipment, supplies, rent and office staffing costs for conducting procedures, which Medicare relies upon to establish payment levels for physicians' services.
[17] For further details, see Getz, Wenger, Campo et al. [2008].
[18] Very similar figures are obtained for the slightly larger set of protocols where the investigator contracts do not specify the number of patients at a site, or from a reweighting of the protocol-level statistics by the number of contracts per protocol.
[19] For an historical overview of hedonic price analysis, see ch. 4 in Berndt [1991]; also see Berndt, Griliches and Rappaport [1995] and Berndt and Rappaport [2001].
[20] See Berndt [1991], ch. 4, and Triplett [2006] for further discussion.
[21] Applications are primarily in modeling health care costs and outcomes. See, for example, Duan [1983], Manning and Mullahy [2001], Manning [1998] and Mullahy [1998].
[22] See Triplett [2006] p.34, footnote 41.
[23] Using log(SWE) does not change the sign or significance of the estimated SWE-related coefficient, or result in material differences in the other estimates.
[24] In all models, tests of the joint null hypothesis that coefficient estimates on SWE, lpatient and the various indicator variables were stable over the two time intervals were decisively rejected.
[25] When pooled over time, the number of observations in the top panel regressions is 207,950 (All), 118,477 (US Only), and 89,473 (Rest of World); for the 1989-1999 (2000-2009) regressions, the corresponding numbers of observations are 125,736 (82,217), 66,246 (52,231) and 59,490 (29,983). Of course the number of observations in the various by phase regressions is smaller, with the smallest number being 2,738 for the 2000-2011 Phase I regressions.
[26] The log transform is used for patients because of the high degress of skewness and wide range of this variable. SWE falls in a much tighter range. No substantial differences in the results were obtained using log(SWE).
[27] These arithmetic AAGRs were more stable than were estimates of compounded AAGRs, since the latter were sensitive to choice of initial and end-year time periods.
[28] SWE, lpatients, and year, therapeutic class and phase indicators are also included in the regressions. Estimated coefficients on SWE and lpatients were similar in magnitude to those obtained in the site-level regressions, and statistically significant at the 1% level.