Comment on “Me-too drugs: is there a problem?”

Joseph A. DiMasi, Ph.D.
Tufts Center for the Study of Drug Development
Tufts University

Professor Hollis’ posting on me-too drugs presents a number of interesting economic points and conjectures. I will offer some thoughts in this message on some of the major points in his posting and related issues.

There is no standard definition of a “me-too” drug. The term has been used rather loosely over a number of decades. Professor Hollis proposes a definition for his analysis. That’s fine. One can restrict analysis to whatever one wants, as long as one is clear about what is being analyzed (although I question below whether the definition he proposes is what he really intends). However, by narrowing focus too much on a theoretical level, one runs the risk that the analysis will focus on constructs whose relevance to the real world is highly questionable. I will address an aspect of Professor Hollis’ submission from that perspective below. First, however, since references are made to our paper on me-toos (DiMasi and Paquette, 2004), I want to be clear that Professor Hollis’ definition does not match the notion that we used in our analysis. It differs in two major ways. Our research concerned approvals of new molecular entities or their biological equivalents by competing firms, not new versions of existing drugs such as isomers (e.g., Nexium®) or new formulations sponsored by the same firm that initially developed the drug for approval. We will examine those sorts of approvals in further work. Secondly, a me-too or follow-on drug in our paper is any entrant other than the first in a class defined by chemical similarity or mechanism of action and major intended uses.

Our definition of the term me-too is what critics of this type of development historically have generally meant. For that reason alone, it is worth examining me-toos defined in this way. Here we are focusing on drugs with different active ingredients (or, for some protein drugs, differences in molecular structure or formulation that have important effects) in the same therapeutic subclass (e.g., different statins or ACE-inhibitors). Criticism has been heaped on development of all follow-on members of these therapeutic subclasses. A common caricature of the process of drug approval is that a breakthrough (first-in-class) drug appears and is marketed successfully for a while before other firms decide to jump in, at little cost and effort, with minor variants of the first-in-class molecule. Our data refutes that notion by providing evidence that development is much more often done more or less in parallel, rather than sequentially. Another objective of our research was to determine trends in the speed with which therapeutic class competition occurs and examine reasons behind the trends. Our discussion of policy proposals and their potential effects on innovation was just that, discussion, not conclusions drawn from an empirical investigation. The same may be noted about Professor Hollis’ discussion of policy, despite his framing a conclusion about innovation incentives as inevitable. That said, I want to be clear about what was suggested.

Although, Professor Hollis appears to divide our discussion of potential problems with a Relman/Angell proposal (Relman and Angell, 2002; Angell, 2004) for imposing costly
regulatory approval hurdles on second and later entrants in a therapeutic subclass into two independent parts, they really were combined in our thoughts and in the paper. We were considering the combined effects of an increased risk of failure, higher costs from more stringent standards, moving targets, and uncertainty about the whole process when we wrote about effects on innovation that “could” occur. It is also important to realize that Relman and Angell did not provide any operational details in making their proposal. So, our comments were based just on what they wrote, interpreting it literally as meaning that sponsors of follow-on drugs would need to prove that their drugs were at least no worse in all drug attributes and superior in at least one attribute. In comparing drugs, there are potentially dozens of side effects that one might consider (although some are more important than others) with differing frequencies of occurrence, a plethora of potential drug-drug interactions, and even multiple efficacy endpoints and indications. Powering randomized controlled clinical trials to provide definitive answers to all comparisons where clinically significant differences might exist and against all competitor drugs in a class could be enormously expensive and often not feasible.

Many of the critics of follow-on drugs simply assert that if we restricted their approval firms would then somehow be “forced” to look for more innovative treatments (Angell, 2004, p. 242). They do not tell us why this would be so. Professor Hollis’ discussion at least is more sophisticated in that he attempts to base his conclusion on economic factors that impact profit maximizing behavior (although, as shown below, his numerical example misses a number of key factors that work against his conclusion even within his highly restricted framework, and he does not provide an empirical justification for his conclusion). It is often easy to be blinded by the models that one examines, particularly simple ones. Professor Hollis’ numerical example is designed to show that innovation would clearly increase with a Relman/Angell-like policy (his policy variant is different in that it uses specific definitions from orphan drug regulations to define “me-tooism” and what it would take for a follow-on drug to obtain regulatory marketing approval). It is a model that makes extreme assumptions. Specifically, the underlying explicit and implicit behavioral and market assumptions for the model are that firms in the current situation seek to develop drugs that are perfect substitutes for one another, there is no price competition among the perfect substitutes (although relaxing this assumption will strengthen his argument about private incentives if demand is inelastic, other things being equal), marketing among competitors can only result in cannibalization of other firms’ sales (i.e., shifts only in market share, with no market expansion possible), and (implicitly, because it is ignored) the policy shift will not result in any changes in risks or expected R&D costs. None of these assumptions is likely true individually, let alone in combination. For that reason alone, the empirical relevance of the analysis is questionable.

Risk of Failure

I will try to remain as close as possible to Professor Hollis’ limiting case assumptions in showing through a numerical example how aspects of the policy change that he did not account for in his numerical example work in the direction opposite to that which he argued. A number of studies have shown that there is near total rent dissipation in the pharmaceutical industry (the literature on ex-post profits show low positive economic profits in recent decades and somewhat negative

1 If such comprehensive testing were actually done, we would likely also be left often without a dominant outcome (i.e., comparisons where, on average, some attributes turned out better and some turned out worse).
economic profits for drugs introduced in the early 1970s). As a simplifying assumption, let us assume that R&D costs and entry in a class will drive economic (not accounting) profits for the class as a whole to zero.

Suppose now that the total profit potential (maximum quasi-rents) of the market for drugs in an existing class is $10 billion in sales net of production and distribution costs on a present discounted value basis (for the moment I will ignore marketing costs). Suppose also that in the current regulatory structure there are five entrants in the class. I will examine expected economic profits in the current situation versus one in which there is a monopoly in the class induced by a different regulatory approval standard. Prior to a demonstration that members of a class will have acceptable risk/benefit ratios there will be an expectation of a considerable risk of failure for the class as a whole. There is generally also a risk of failure for drugs in classes even where some members of the class have acceptable risk/benefit ratios. For simplicity, I will ignore this latter risk and consider only an approval success rate for the class. Suppose that prior to entry potential entrants to the class perceived that there was a 25% chance that drugs from the class would meet current regulatory approval standards, and also that all five entrants share the market equally (I am abstracting here from complicating factors such as differences in when development is initiated, the length of that development, differing out-of-pocket costs of development, and order of entry effects in the marketplace). Given these assumptions about economic profits, technical risk rates, and market share, the implied expected cost of development per firm is $500 million on a present discounted value basis.

Now assume a regulatory policy change that ensures that only one drug in the class will be approved (assuming that other members of the class are not tested and proven to be “better” than the one approved drug in the class). Professor Hollis maintains that innovation incentives are increased in this scenario. Presumably that means that incentives are greater even in classes that otherwise would have had entrants in the current system, so that there will be at least as many firms with drugs in such a class that would develop them for regulatory approval as there are under the current system. To be conservative, let us assume that the number of firms that will try to develop drugs to attain the monopoly position is the same (five) as the number of firms that developed drugs in the class under the current system. We will assume that each of these firms is equally likely to win the race for the induced monopoly position. Under the assumptions about technical approval success rates that we have made, each firm will view its prospects as a 5% chance of earning $10 billion in net sales at an R&D cost of $500 million and a 95% chance of simply losing $500 million in R&D costs (assuming full development for all members of the class). The expected economic profits per firm in this case would also be zero. Thus, despite earning much higher sales if a firm succeeds in becoming a monopolist, incentives are not increased under the new regulatory policy because of the increased risk of failure.

2 In the context of this simple model, this result is completely general. Let \( p \) be the subjective probability, shared by all potential entrants, that drugs in the class will have acceptable risk/benefit ratios. Also, let \( s \) be the equal market share achieved by entrants should drugs in the class make it to the marketplace. Finally, let \( V \) be the aggregate present discounted value of all future net sales and let \( R \) be the R&D cost that firms must incur if they decide to try to get a drug in the class approved for marketing. The expected economic profit per firm is then \( p(sV-R)-(1-p) \cdot R \). Setting expected economic profits equal to zero and solving for \( R \), yields \( R = p \cdot s \cdot V \). In the induced monopoly scenario, expected economic profits per firm are \( p \cdot s \cdot (V-R)-(1-p \cdot s) \cdot R \). Substituting the value for R&D costs per firm determined for the current scenario into the expected economic profits expression for the induced monopoly scenario also yields a value of zero.
The reduced likelihood of reaching the market at all in the induced monopoly scenario is a factor that is ignored in Professor Hollis’ model. It may also be the case that firms are not risk-neutral over all sets of potential risk/reward outcomes, but that they could become risk averse if the distribution of risks across outcomes becomes skewed enough. That is, firms may not be willing to take a “fair bet” or even a somewhat better than fair bet if the there is a sufficiently high risk of a substantial loss. This is one reason why innovation incentives could in fact be diminished under the regulatory policy change.

**Marketing Costs**

Professor Hollis essentially offers just two reasons to justify his conclusion that innovation incentives will increase with the assumed policy change. One is that firms in the current regulatory environment have to share markets, as opposed to capturing the entire market in the induced monopoly scenario. As we have seen, this factor is offset by the increased risk of incurring R&D costs and failing to reach the marketplace in the induced monopoly case. Professor Hollis’ other reason for concluding that innovation incentives increase is that aggregate marketing expenses would be greater in the multiple entry scenario compared to an induced monopoly scenario with no market expansion. However, the assumption of no market expansion seems an unlikely outcome for real world settings. For example, in a recent study of the effects of direct-to-consumer (DTC) advertising done for the Kaiser Family Foundation, Rosenthal et al. found that $1 of DTC advertising was associated with a $4.20 increase in sales at the class level, but that no effects on market share from increased DTC spending at the product level could be detected from their regressions (Rosenthal, et al., 2003). Similarly, their product level regressions showed no statistically significant market share effects for detailing expenditures.

Other studies have also shown market expansion effects and not shown market share effects for DTC advertising, or market share effects only in a portion of the market. Some of the literature does find market share effects for detailing, but also finds market expansion effects with increases in total class marketing expenditures. In light of the empirical evidence, it is very hard to maintain the position that the marketing efforts of additional rivals in a class will merely affect market shares, with no expansion of the total market.

Aside from noting the empirical evidence, we can point to a number of subtle reasons why markets effectively expand with rivalry. Inventions do not immediately diffuse fully. It would take the monopolist some time to capture the full quasi-rent potential of a market. If there are a number of competing firms in the marketplace, the movement toward the capture of maximum quasi-rents will likely be much quicker. Thus, the present discounted value of aggregate quasi-rents will be larger in the rivalrous than in the monopoly scenarios. In this way, we can say that the market will expand with rivalry. Furthermore, the market will expand if additional indications are approved, or at least approved more rapidly, under rivalry.3 It is also thought that compliance rates are poor for many drugs (even for those with insurance coverage). Marketing in the rivalrous scenario could improve compliance rates in general for the class. Finally, as long

---

3 The pursuit of different indications by rivals can occur prior to original marketing approval, as well as later. For example, although Celebrex® and Vioxx® were approved just five months apart, Celebrex® was originally approved for osteoarthritis and rheumatoid arthritis, while Vioxx® was originally approved for osteoarthritis, acute pain, and dysmenorrhea.
as there is some elasticity to demand, price competition in the rivalrous case will lead to some market expansion (see the discussion below on price competition). Market expansion may or may not be good from a societal perspective (it depends largely on whether drugs would otherwise be under- or overutilized in each of their uses), but it increases private incentives, and so works against Professor Hollis’ argument.

Even if we put aside market expansion effects and assume that the total market size will always be the same for both the current and the induced monopoly systems, Professor Hollis argues that aggregate marketing expenditures would be greater under the current system. That may be so under a no market expansion assumption, but how much more would be the relevant issue. Marketing expenditures will not be much greater in the rivalrous marketing scenario to the extent that marketing costs are proportionate to sales. In fact, if the marketing-to-sales ratio is constant regardless of market structure, then there would be no difference in aggregate marketing expenditures as between the current and the induced monopoly systems, and so no increased incentives to innovate under the induced monopoly scenario in Professor Hollis’ model.

There are, however, theoretically at least two possible reasons why there could be higher aggregate marketing expenditures in the rivalrous case. The general theoretical literature on advertising presents the possibility that for some types of goods (so-called experience goods) marketing intensity levels may be higher for later entrants. This, however, is usually explained as the need to inform consumers who have had satisfactory experiences with existing products about the qualities of one’s new product to get them to at least try the new product. The presumption is that the products have different attributes. The second factor is the impact that economies of scale in marketing might have. However, it is not clear in this context that there would be substantial economies in serving a market of a given size with one firm as opposed to with a small number of similar firms. Given that we are discussing what would presumably be a long-term policy change and that we are analyzing pre-innovation conditions for a class, economies of scale must also be considered here in a long run context (where infrastructures can be adjusted to better fit a new regulatory paradigm). It is also the case that the marketing infrastructures would be covering drugs that lie outside of this framework (drugs that have advantages and drugs that are unique with regard to chemical structure or mechanism of action), thereby spreading those costs over more drugs than are the subject of analysis here.

Finally, any decrease in marketing expenditures for the induced monopoly scenario would have to be weighted by the likelihood that any potential entrant would achieve them. Suppose that the $10 billion noted above is what the five firms earn in the current scenario in the aggregate after marketing, but not R&D, costs are considered. If we have lower marketing costs for the induced monopoly scenario, then the amount earned by the monopolist in net sales inclusive of marketing costs would be higher by this amount. However, when considering pre-innovation expected profits per firm, the lower marketing costs should be weighted by the probability both that drugs in the class will have acceptable risk/benefit ratios and that any particular one of the five firms would be the one to attain the monopoly position. In the numerical example given above, that probability would be just 5%. In sum, increases in marketing expenditures under rivalry may

---

4 If we use the same terminology from in footnote 2, except that we redefine V to be aggregate sales under the current regulatory scheme net of aggregate marketing as well as aggregate production and distribution costs, then we again have $R = p\cdot s\cdot V$. Now, let $m$ be the present discounted value of the amount by which aggregate marketing...
have little impact on the expected profit calculations, and, in theory, could be more than offset by market expansion.

As an aside, I would note that much of the criticism of follow-on drugs flows from assumptions about the nature of the effects that marketing has on demand for these products. The argument is often made that these drugs do not really differ much from the first entrant, but marketing is used to create perceived, but not real, differences in the minds of physicians and patients. In other words, marketing efforts increase product differentiation beyond that which would exist if consumers and/or their agents (physicians) had been given the most impartial and comprehensive information that existed at the time. However, to the extent that this is true, it can be largely remedied, as we suggest in our paper, by improved information on drugs already in the marketplace (e.g., pharmacoepidemiology studies, disease management programs, pharmacoeconomic analyses, so-called pragmatic clinical trials, and, occasionally, additional randomized controlled clinical trials [not necessarily conducted by manufacturers]).

Moving Targets and R&D Costs

A cost to the proposed regulatory change that we noted in our paper, but have not yet considered here, is the impact on expected R&D costs that the regulatory change would engender, from both increased approval standards and moving targets. These are costs (potentially quite substantial) from the regulatory change that would have to be added to the costs noted above that arise from a higher risk of sinking large expenditures into R&D programs of the type that exist under the current system that will now not yield any returns. Here, we need to also consider that the more stringent approval standards will result in higher expected development costs for the average program, even if the approval targets are known in advance (i.e., which drugs will serve as comparators). This is so because there is some likelihood that firms that lose the race to be first will shift focus and spend additional funds to try to demonstrate superiority. This strategy may be perceived to be a worthwhile risk because at that point in time much of the firm’s R&D expenditures to develop the drug are sunk costs. In addition to higher expected development costs from testing to show superiority against a known comparator, the expected costs for firms considering entry will be still higher because there is a likelihood that part of the development program would have to be stopped and begun again as other drugs reach the marketplace (i.e., firms would be faced with the prospect of a moving target). This latter scenario may be replicated a number of times during development. Aside from directly increasing R&D costs, this will delay the introduction of follow-on drugs. Delays in entry can substantially reduce expected profits.\(^5\) All of these factors serve as strong disincentives to development when viewed from a pre-development perspective.

\[\text{expected profits for the current regulatory environment would exceed those that would be incurred by the monopolist under the induced monopoly regulatory scheme. Then, expected profits per firm with the new regulatory scheme become } p \cdot s \cdot (V + m - R) - (1 - p \cdot s) \cdot R = p \cdot s \cdot m. \text{ If marketing-to-sales ratios remain the same in the aggregate regardless of market structure, then } m = 0 \text{ and expected profits are zero in the induced monopoly scenario. In that case, marketing would have no effect on incentives. Otherwise, if aggregate marketing expenditures decline for the induced monopoly case, expected profits will vary proportionately with the likelihood of reaching the marketplace first with a drug with an acceptable risk/benefit ratio.}\]

\(^5\) Of course, from a pre-innovation perspective, the effects of delays must be weighted by the likelihood that they will be incurred.
Professor Hollis recognizes the moving target problem and proposes as a solution, restricting the drugs to be used as comparators to those approved prior to an 18-month window leading up to regulatory submission of an application for marketing approval of an investigational drug. This solution appears completely arbitrary. Or, at least, one can say that Professor Hollis has not provided any rationale for why 18 months is the right figure to “largely solve the problem of the moving target.” What we can note, however, is that development time data indicate that phase III testing has averaged about three years. This is an average, so there are drugs for which phase III testing has lasted longer than three years. In addition, arguably, at least some phase II trials might be affected by whether testing against comparator drugs is to be undertaken, for what purposes, and which comparators will be used. So, should we make the window say, five years, seven years, eight years, etc.? What would be left to test against with such lengthy windows for narrowly defined classes? Our data showed entry from therapeutic substitutes in the same class occurring relatively rapidly for recent periods. I would also note that while firms do project timelines for the development of their drugs, early projections often turn out to underestimate the subsequent reality as a consequence of a variety of unexpected circumstances.

Investigational drugs that would have to be tested against comparator drugs in the same class with windows this lengthy would be very late entrants to the class. It is highly likely that firms with such late entrants would not bother pursuing entry unless they already had reason to suspect that they could demonstrate superiority for some attributes (even if this is done post-approval) or could find some niche that is not already well served. One would have to wonder, then, if a proposal of the type suggested by Professor Hollis with a window lengthy enough to “largely solve the problem of a moving target” is a solution for which there is no problem.

Innovation Races and Product Differentiation

An underlying assumption in Professor Hollis’ framework for analyzing innovation incentives discussed above is that firms in the current environment are willing to compete consciously in the development of perfect substitutes. However, pharmaceutical company executives routinely comment that they are not interested in developing what are effectively identical products. Observations of how drug development tends to proceed appear to be consistent with these characterizations. As Scherer has noted (Scherer, 1996): “If there is no future in me-too drugs, what companies do is to seek an unpopulated niche. Given the long lead times in drug discovery and testing, however, they may inadvertently end up with a me-too drug. But they are not trying consciously to duplicate the output of one or more known racing partners.” Scherer (1996) further describes a type of racing that characterizes drug discovery and development as follows: “Drug discovery and development are more like a marathon than a 400-meter dash. In a marathon, the runner mostly has to pace himself, or he will drop from exhaustion. But there may be key stages at which he comes into head-to-head rivalry. Moving from that analogy to a closer approximation, firms in the pharmaceutical industry are racing to horizontally differentiated niches in what economists call product characteristics space. If that is true, and I believe it is, it is not an “either he wins or I win” race. It can be an “everybody wins” race – one, to be sure, in which some win more than others.”

The product differentiation model just described increases the complexity of the entry calculus discussed above, although it does not eliminate the costs and benefits of the regulatory change
described above as factors that affect incentives. However, it further weakens any basis for Professor Hollis’ claim that the change in regulatory policy would clearly increase incentives for private drug development. In making their strategic decisions in the current system, firms will be looking at the prospects for some market expansion from differentiated products (e.g., in terms of new patient groups, new indications, and patients that have tried a competitor product but did not respond to it or who suffered side effects from it), not simply splitting a fixed market with identical products. The net impact of all the factors that we have considered from the proposed regulatory change, including changes in marketing intensities and price competition, is ultimately an empirical question, but it seems very unlikely to me that on average the direction of effect will be increased development incentives. Rather, the reverse is likely true.

Creative Destruction

Professor Hollis notes the Lichtenberg and Philipson (2002) paper (which we also do in our discussion of factors that might affect the speed of entry) and uses it to support his conclusion that me-too drugs reduce incentives for innovation because they take market share from pioneers (p. 3). However, we have already discussed above how, from a pre-innovation perspective, a larger market from winning a winner-take-all race is offset by a higher risk of failing to win the race. It is also worth noting that the data used in the Lichtenberg and Philipson (2002) paper were quite broad in comparison to the kinds of groupings we used for our paper. We noted in our paper that the classes we used were conservative from a full economic perspective. Drugs compete in the marketplace across, as well as within, classes of the type we analyzed. It was not our purpose to characterize and analyze demand in complete markets for drugs; our focus was on evidence on the development and entry of drugs in what many would think of as me-too groupings. Professor Hollis recognizes that the empirical analysis in Lichtenberg and Philipson (2002) used very broad categories, but then he effectively dismisses the distinction. He essentially argues that nearly all of the lost value for new drugs comes from later entry in the same class. In that regard, it is instructive to note just how much broader the classes are in Lichtenberg and Philipson (2002) than anything we are discussing here. For our paper, the mean number of entrants (including the first-in-class) was four and the median was three. In the Lichtenberg and Philipson (2002) analysis, at the time the average new entrant was launched the mean number of drugs in the class that were already on the market was 25.

Lichtenberg and Philipson (2002, p.644) have firmly in mind the kind of creative destruction of value that comes from improved products (“…the demand for a given innovation is often destroyed by the entry of new, superior products long before patent expiration”). Additional competition after entry in their empirical analysis can come from drugs in newer classes, not just new entrants to the same class. In theory, if the advances are large enough, new and improved classes can destroy more value than additional entrants to an older class.

The impact of generic competition in Lichtenberg and Philipson (2002) is limited to a significant extent because of the fact that generic competition occurs much further out in time than does therapeutic class competition. Thus, on a present discounted value basis, the effects of generic competition are greatly diminished. Professor Hollis appears to not realize this point in his discussion of the benefits of generic and me-too drugs (pp. 3-4). He notes that it has been estimated that drug costs were lower by $8 to $10 billion in the United States in 1994 because of generic competition. He then assumes, without empirical support, that the gain to consumers
from lower generic prices presently is $40 billion and that the benefits of me-too drugs do not match that figure. Whatever the current real values are for generics and follow-on drugs, this discussion misses the fact that generic competition occurs later in time. Thus, any comparison would have to either capitalize the benefits of follow-on drugs forward or discount the benefits of generic drugs back in time to when follow-on competition occurs (or some other common point in time). Therefore, a comparison of the type that Professor Hollis’ uses (albeit his is one in which he does not actually have a specific value for the benefits of follow-on drugs) grossly overestimates the benefits of generics relative to follow-on drugs. It is also unclear what significance there is to such a comparison.

**Price Competition**

Professor Hollis cites two studies as the basis for his claim that me-too drugs engender very little price competition (Lu and Comanor, 1998; Ekelund and Persson, 2003). Specifically, Professor Hollis cites results in the Lu and Comanor (1998) paper on average relative launch prices for drugs rated C by the FDA (under the old three-tiered therapeutic significance rating system). A number of points can be made about this paper to dispute Professor Hollis’ conclusion that there is nearly no price competition from the introduction of new drugs in a class.

First, the data used for the Comanor and Lu (2003) paper are now quite old. They refer to new drugs introduced from 1978 to 1987, a period during which price discounting was limited relative to what we see today. As the authors themselves note (p. 133): “The period covered in this study largely predates the rapid growth of managed care, which is relevant because it led to much greater price discounting. For this reason, the results presented here may not fully represent current conditions in pharmaceutical markets.” On the other hand, we discussed in our paper results from a study that I conducted on pricing trends for drugs approved from 1995 to 1999 (DiMasi, 2000) that showed an average launch price discount of 14% for new entrants relative to the weighted mean price of existing drugs in the class. These results do not even include the impact of rebates (which were unavailable).

Second, relative launch prices, as measured by Lu and Comanor (1998), were often higher than existing prices for drugs introduced during their study period. The average relative launch price was over 3 for drugs rated A by the FDA and over 2 for drugs rated B by the FDA. The impact of additional entrants to a class, however, can be measured in such a period by careful statistical analysis. Lu and Comanor (1998) ran a number of regressions, controlling for various factors, to examine the extent of competitiveness for this period. Contrary to Professor Hollis’ conclusion, the authors themselves concluded that entry induces a significant amount of price competition. As they note in their abstract: “In addition, the number of branded substitutes has a substantial negative effect on launch prices, which reflects the importance of competitive pressures.” For a quantitative assessment their regression results imply, as they note (pp. 114-115), that: “increasing the number of direct substitutes from one to two leads on average to a 38% decline in the ratio of a product’s launch price to the average price of its predecessors; while increasing the number of substitutes from two to three leads to a 19% decline.” Similarly, they also evaluate through regression analysis the impact of entry on pricing dynamics and concluded the following (p. 116): “More numerous rivals have the expected effect of slowing price increases.” In sum, a proper reading of the Lu and Comanor (1998) study strongly supports the hypothesis that
therapeutic class competition in a relatively unregulated market induces a significant amount of price competition.\(^6\)

The Ekelund and Perrson (2003) study was designed to examine price competitiveness in a price regulated market. They essentially replicated the Lu and Comanor (1998) analysis for drugs launched in Sweden from 1987 to 1997. Their regression results show that the impact of additional entrants on launch prices has the expected sign (negative) but that the effect is not statistically significant. The message from the Ekelund and Perrson (2003) study is not that drugs in a therapeutic class cannot compete on price. The message (to the extent that one can embrace a null hypothesis) is that price regulation suppresses price competition.

Misuse of FDA Statistics

Given the frequent use by industry critics of the FDA’s therapeutic significance rating system for new drug product approvals as a de facto indicator of whether a new drug adds any value, it is worth examining the propriety of such gross generalizations. The FDA’s rating system is a management tool designed to improve its internal allocation of resources. It would not be an effective management tool if all or nearly all drugs were rated the same way. To some degree it must therefore be endogenous. Also, what constitutes a significant difference is to some extent a judgment call, the FDA will typically make such determinations only on the basis of randomized controlled trial evidence, and whether there is a strong medical need that is not met or is fairly poorly met by existing therapies will play a substantial role in assigning a priority rating.

The FDA maintains that its ratings at the time of approval are not intended to predict a drug’s ultimate value. In addition, as far as I know, firms have not used the fact that their drugs had received priority ratings to promote those drugs. Furthermore, testing to meet regulatory standards is designed only to show average effects. That is, what is examined is whether the mean effect of a drug is statistically significantly different from placebo or an active comparator. If drug effects are idiosyncratic, then average treatment effects will not account for the full therapeutic value of a new drug.\(^7\) For example, if we were to show that two different drugs were equally effective in treating 40% of the population using standard testing methodologies, we still would not know if it is the same 40% for both drugs.\(^8\) The testing, and so the ratings, also would

---

\(^6\) This is not to say that the entry of therapeutic class substitutes will induce the type of cutthroat price competition seen with generic entry, where products compete essentially just on price. Branded drugs will compete on quality attributes as well, a notion that is also supported by data in Lu and Comanor (1998) and DiMasi (2000); the results on pricing by therapeutic rating at least suggest that prices adjust to some extent to quality differences.

\(^7\) This holds whether comparisons are made with active comparators or placebo. See Malani and Hu (2004) for an examination of how value resulting from heterogeneity in responses to different treatments is not captured by the standard testing methodologies used for marketing approval.

\(^8\) Angell (2004, p.90) maintains that it would be “easy” for a drug company to test whether its drug works for some patients who have failed in some way on another drug in the class. While it may be conceptually easy to establish the methodology for this type of testing, it would often not be logistically easy or financially viable to do so in pre-market testing. If this were to be done universally, in most cases firms would have to run trials with patients given the comparator drug first, and then its own drug for the patients who do not respond well to the comparator drug. Regulators would also likely want to see the results of testing the firm’s drug against placebo (if ethical) to better establish effect sizes. The firm would also want to test against placebo in the general population (i.e., different additional trials) since it would not want a label that restricted the approved uses of its drug to just second or later-line therapy when first-line therapy appears to be a technically and financially feasible option. The number of
not reflect full added therapeutic value even in cases where one is not able to conclude that a new
drug is better from some overall assessment of attributes for the patient population as a whole,
but where the new drug nonetheless offers trade-offs in attributes relative to existing drugs that
different patients value differently.\textsuperscript{9} It is therefore wrong to conclude that a non-priority
(standard) rating from the FDA indicates that there are necessarily no meaningful differences
between such a drug and others already on the market.

It is also important to note that the FDA and other regulatory authorities do not have a
presumption of identical effects for all drugs in the same class. Thus, when one drug has its label
expanded to cover a new use or patient subgroup, the regulatory authorities will not assume that
all of the other drugs in the class should also be so labeled. Or, when a safety issue is noted for
one drug in a class that is serious enough to prompt its withdrawal from the market, the
regulatory authorities will not automatically move to withdraw all other drugs in the class. For
example, when Baycol\textsuperscript{®} (cerivastatin) was removed from the market after reports of an
unusually high incidence of an uncommon serious side effect (rhabdomyolysis), responsible
observers and the regulatory authorities did not maintain that all statins should be taken off the
market. A particular drug effect (good or bad) may be unique to only one of the marketed drugs
in a class, or the effect may be a “class effect.” However, even for class effects, the extent to
which the effect is realized can vary significantly by drug. For example, rhabdomyolysis is
thought to be a class effect for statins, but the incidence of this side effect for other approved
statins appears to be much lower than for cerivastatin.\textsuperscript{10}

Professor Hollis does note and rely on an analysis in Love (2003) that uses the FDA therapeutic
ratings for new molecular entities (NMEs) approved from 1993 through 2002 to claim that “the
share of investments in new products that have significant improvements over existing
treatments is 20 percent, with 80 percent of the investment in new products spent on projects that
demonstrate no significant improvement over marketed products.” This percentage breakdown
is based only on two ratios. There are a number of problems with the calculation. One of the
ratios used is the percentage share of NMEs that were priority-rated by the FDA for NMEs
approved from 1993 to 2002. The value of the ratio is based on an arithmetic error. Love (2003,
patients required for clinical trials would therefore expand substantially. Thus, clinical trial costs would balloon
and, even if positive effects were found with the investigational drug for poor responders or nonresponders to
existing drugs, little or no additional practical information (beyond that from approval as first-line therapy) would be
obtained as a result. The reason that the additional knowledge obtained would not generally be very useful is that,
few markers have been developed and implemented to, on an ex ante basis, identify individual patients who would
be poor responders or nonresponders to one drug as opposed to another. Physicians would generally then still, as
they do now, have to use the drugs on a trial and error basis to find one that works well for an individual patient.
\textsuperscript{9} That is, the new drug is not better, just different. For example, testing may indicate that the new drug has a
reduced potential for some side effects, but increased potential for others. Or, there might be a trade-off between an
increased potential for greater efficacy and an increased likelihood of side effects. Then, the introduction of the new
drug and the greater product diversity that it brings may result in a better matching of patient preferences and drug
attributes.
\textsuperscript{10} For example, in an observational study of the incidence of rhabdomyolysis for lipid-lowering drugs, Graham et al.
(2004) found incidence rates for this serious side effect that were 12 times higher for cerivastatin than the average
for atorvastatin, pravastatin, and simvastatin. Similarly, reports of the deliberations of the recent FDA Advisory
Committee meeting on COX-2 inhibitors indicate that committee members concluded that cardiovascular risks are a
class effect, but that the risks are much lower at common treatment doses for Celebrex\textsuperscript{®} than for Bextra\textsuperscript{®} or
Vioxx\textsuperscript{®}.
p.17) has a table giving the annual numbers of priority and standard NMEs for 1993 to 2002. The annual numbers are correct and the sum of standard NME approvals is correct. However, the sum of priority NMEs given in the table is wrong. Love (2003) has 79 priority NMEs in total for this period when the actual number is 120. The correct number raises the percentage of NMEs with priority ratings from the 31% that Love (2003) has to 40%.11

The other ratio used is a ratio of clinical trial sizes for priority and standard-rated drugs, where the clinical trial sizes were obtained from some FDA documents. These documents do not include discussions and data for all clinical studies conducted on investigational drugs. This is acknowledged in Love (2003), but it is argued there that this should not affect the ratio of clinical trial sizes. That is somewhat dubious given that the differences between clinical trial sizes obtained in this way and actual clinical trial sizes are likely highly variable by drug. However, even if we put that concern aside, there is no reason to believe that the amounts spent during the clinical testing period on a per subject basis is fixed across therapeutic ratings. This is an implicit assumption in the Love (2003) calculation. Results in DiMasi et al. (2003) suggest that costs per subject for priority-rated approved new drugs were 16% higher than for standard-rated drugs (down from 42% higher for an earlier period with a three-tiered rating system [DiMasi et al., 1991] with the two higher ratings compressed into one group). If we applied the correct share of NME approvals for priority drugs for the period given and a 16% higher cost per subject for priority NMEs, then the expenditure share for priority NMEs increases from 20% to 44%.

The estimation of such shares for approved drugs, however, is of questionable validity and significance. Two important factors not considered in these calculations are discovery costs and clinical success rates. If discovery costs are higher and success rates are lower for priority drugs, then an estimation of the type in Love (2003) could substantially underestimate the share of total expenditures attributed to priority drugs. Beyond these considerations, attempts to draw normative conclusions based just on distinctions between spending on priority and standard drugs is highly questionable given results in DiMasi and Paquette (2004) that indicate that development of drugs that eventually get standard ratings is often done more or less in parallel with drugs that get priority ratings and because the likelihood of finding at least one molecule in a class that has an acceptable risk/benefit ratio increases with the number of independent development efforts.12

11 Prior to enactment of the user fee program and the priority/standard two-tiered rating system, the share of NMEs that were rated the equivalent of priority was 49% (A and B ratings in a three-tiered system). For the two years subsequent to the end of the period used in Love (2003), the share of NMEs that received priority ratings increased to 43% for 2003 and 56% for 2004 (including five biologics moved to FDA’s CDER from CBER in a reorganization in 2003 that shifted the review of most biologics to the drug center; the share was 52% excluding the five biologics).

12 See, for example, discussions in Scherer (1966) and Scotchmer (2004) regarding innovation in general. To put the matter formally in a simple context, suppose that N firms pursue different independent approaches to developing a drug in a new, but unproven, class, and that the probability of failure, q, is the same for all firms. Then, the probability that at least one investigational drug will be found that has an acceptable risk/benefit ratio is 1-q^N. Thus, the likelihood of at least one success increases with the number of firms pursuing different molecules in a class. This observation alone does not tell one what the socially optimal number of experimental approaches is, but we would certainly expect that in general it is greater than one.
Placebo Control in Practice

Although Professor Hollis did not address in any detail the nature of current clinical testing practices, it is worth noting a few realities. The impression that one often gets from superficial critical discussions of drug industry clinical testing programs is that, except perhaps on rare occasions, drug companies do not test their drugs against anything other than placebo. Nothing could be further from the truth. There are a substantial number of trials where investigational drugs are tested against standard treatment because the use of placebos is thought in those cases to be unethical (e.g., cancer and AIDS trials). Drug companies may also, for strategic reasons, sometimes test their drugs against active comparators (either pre- or post-original approval). The motivation for firms to conduct comparative testing will likely increase over time as the marketplace demonstrates a demand for such information. Finally, testing against active comparators has often been done to satisfy European regulators. Although drug reviews are considered on a case-by-case basis, EMEA guidelines (European Medicines Agency, 2004) indicate that in cases where the use of a placebo is considered ethical and existing medicines are available to treat a condition, pivotal clinical trials should have three arms (investigational drug, placebo, and active comparator). The standard for testing against the active comparator, however, is noninferiority, not superiority, for some major attribute. Approval of drugs that test better than placebo but worse than the comparator for the major attribute may still be approved after consideration is given to all aspects of quality, safety, and efficacy.

Lessons from Orphan Drug Regulation?

Professor Hollis points to US orphan drug regulations and their perceived impact on the development of drugs for orphan indications to both provide a framework for defining me-too drugs and to argue that the experience with orphan drug regulation suggests that therapeutic class competition should be precluded along the lines discussed above. The US orphan drug legislation affords a number of incentives for orphan drug development, including tax credits, clinical trial grants, advice from regulators, waiver of user fees in some instances, and marketing exclusivity. This last factor is usually thought to be the most powerful of the incentives. It is important, however, to understand just what problems this solution addresses.

During the period of marketing exclusivity, orphan drug regulations preclude approval of other drug products that are the “same drug for the same indication.” Professor Hollis suggests using the definition of “sameness” in the orphan drug regulations to define which drugs should be grouped and termed me-too drugs. With all due respect, Professor Hollis appears not to understand the concept. He lists the drug products Crestor®, Clarinex®, Levitra®, and Lexapro® as possible examples of how the notion of sameness under orphan drug regulations can be extended to define me-too drugs for drugs in general. These are all small molecule NMEs. They contain active ingredients that are different than those contained in the first-in-class approvals for their respective therapeutic classes (or for any other drugs in their classes, as we have defined therapeutic class in our paper). Under orphan drug regulations, for a small molecule product to be considered the “same” as another drug product that has received orphan drug marketing exclusivity, it must contain the same active ingredient. Therefore, for small

13 For purposes of this discussion I will refer to both drugs and biologics as drugs.
molecule drugs the orphan drug regulations do not prevent the type of therapeutic class
competition that Professor Hollis seeks to prohibit.\textsuperscript{14}

The issues for macromolecules under the orphan drug regulations are a little more complex. It
was about 10 years into implementation of the Orphan Drug Act before the FDA established
rules that provided industry with clear guidance on what constitutes “sameness.”\textsuperscript{15} The delay led
to some high profile conflicts in the biotechnology sector over whether variation in particular
characteristics of the physical/chemical structure of a macromolecule yielded a “different”
drug.\textsuperscript{16} In late 1992 the FDA provided guidance on what constituted sameness for
macromolecules. The presumption under the rules is that macromolecules with small changes in
amino acid sequences, glycosylation patterns, and a number of other small differences in
structure for a variety of types of macromolecules do not have any clinical effect and so the
molecule would be considered the “same” as the drug with exclusivity. Such follow-on
macromolecules could, however, be considered different drugs if the manufacturer of the later
drug demonstrates “clinical superiority” for its product.\textsuperscript{17}

The rationale for orphan drug marketing exclusivity has to do principally with providing
protection against free-riding on research for drugs that were unpatentable (a particular concern
for biopharmaceuticals) or that had patents that had expired or would expire prior to regulatory
approval or shortly thereafter. The regulations on “sameness” effectively protect firms against
early generic or generic-like\textsuperscript{18} (in the case of macromolecules) competition, not therapeutic class
competition of the type that we have discussed above.\textsuperscript{19}

Moving beyond “sameness,” it is unclear to me whether the concept of “clinical superiority” in
the context of the orphan drug regulations properly understood is what Professor Hollis really
had in mind for an approval hurdle. It certainly does not seem to be what Relman and Angell
had in mind. It is worth expounding briefly, then, on just what the concept means. Under the

\textsuperscript{14} Although the share of new drug orphan approvals that are for small molecules has trended downward, small
molecule drugs still constitute a majority of these approvals over the entire history of the implementation of the
Orphan Drug Act.
\textsuperscript{15} See Shulman and Manocchia (1997) and Rohde (2000) for discussions of major developments in U.S. orphan drug
regulation.
\textsuperscript{16} In a dispute between Genentech and Lilly about their versions of human growth hormone, the FDA ruled that
Lilly’s product was not the “same” drug as Genentech’s earlier approved product, even though Lilly’s product
differed structurally only by the addition of a single amino acid to a sequence of 191 amino acids. The ruling made
Lilly’s product eligible for approval and orphan drug exclusivity.
\textsuperscript{17} These differences in large molecular structures might be viewed as something akin to differences in excipients in
small molecule formulations. They are not expected to have clinical effects, but in highly uncommon cases they
might. Thus, without evidence to the contrary, these molecules can be viewed as essentially generic versions of
existing drugs. In theory, clinically different outcomes might result not from these structural variations, but rather
from differences in formulation, route of administration, or frequency of administration.
\textsuperscript{18} In the absence of regulations precluding the approval of drugs that are the “same,” firms seeking marketing
approval for macromolecules that are the “same” as products already on the market would avoid some discovery and
other costs, but some clinical trial testing would still be necessary. The situation is similar to that which was
required in the pre-Hatch/Waxman era to obtain approval of a generic drug.
\textsuperscript{19} The protection is somewhat weaker than patent protection on drug substances. Even with orphan drug marketing
exclusivity, it is possible for other firms to get the same drug substance approved during the exclusivity period for
other uses (either nonorphan or other orphan indications) if there are no other relevant intellectual property
protections for the original manufacturer.
regulations, clinical superiority means that the manufacturer has demonstrated a significant therapeutic advantage with regard to some aspect of efficacy or safety, or in unusual cases by showing that the product “makes a major contribution to patient care.” It does not mean that the new product is “better” than the original product in some overall sense for all patients or even on average.

It is worth examining the clinical superiority notion in the context of real examples. As far as I know, only a few orphan drug approvals have been granted on the basis of a demonstration of clinical superiority. Two such cases exist for interferon beta products in the treatment of relapsing-remitting forms of multiple sclerosis (MS). Avonex® (interferon beta-1a) was able to break the exclusivity of Betaseron® (interferon beta-1b) on the basis of a better safety profile when tested against placebo. Rebif® (interferon beta-1a), on the other hand, was able to break the exclusivity of Avonex® on the basis of improved efficacy in a head-to-head comparison.\(^\text{20}\)

The demonstrated efficacy improvement for Rebif® was a short-term increase in the percentage of patients who were free of relapses. However, Rebif® must be administered more frequently, clinical trial data suggested that Rebif® had a worse side effect profile (injection site necrosis, elevated liver function tests, and leucopenia) than Avonex®, and the incidence of the development of antibodies that neutralize the drug’s effect appeared to be greater for Rebif® than Avonex®. Despite the “clinical superiority” label, it may be hard to argue in some overall sense that Rebif® is a better drug than Avonex®. We can certainly say, though, that it is different. In essence, there is a trade-off between an increased likelihood of a better short-term efficacy outcome and an increased likelihood of certain side effects (as well as more ease of use). This diversity in outcomes offers choices for individual patients who, together with their physicians, can determine what is best for them.

Two other points about this case should also be noted. First, despite obtaining a clinical superiority assessment, Rebif® did not receive a priority rating from the FDA. This example illustrates the inappropriateness of using the FDA’s therapeutic rating system to conclude that new drugs that receive a standard rating add no value. Second, as is the case for small molecules, the orphan drug regulations do not preclude other molecules from being approved during the period of marketing exclusivity for the same uses. Copaxone® (glatiramer acetate) is not an interferon-beta, but it was approved for marketing for relapsing-remitting MS during the Betaseron® and Avonex® exclusivity periods. It had also received a standard rating and an orphan drug designation from the FDA.\(^\text{21}\)

In sum, although Professor Hollis points to the orphan drug regulations and their definitions of “sameness” and “clinical superiority” as a model for how to keep me-too drugs off the market unless they are shown in pre-market testing to be better than what is already available, a proper examination of the regulations shows that they would do no such thing. Furthermore, the

\(^{20}\) By the time that Rebif® was approved, the exclusivity on Betaseron® had run out.

\(^{21}\) Recently (November, 2004), the FDA approved an additional late entrant in this therapeutic area. Tysabri® (natalizumab) received a priority rating. Like many late entrants, it was developed with the hope that it would provide certain advantages. Pre-market testing indicated that it substantially reduced the frequency of clinical exacerbations for patients with relapsing forms of MS. It is also infused only once a month. However, as of this writing, the manufacturer has suspended marketing and clinical trial use of Tysabri® while unexpected safety issues are investigated.
apparent success of the Orphan Drug Act in stimulating the development of drugs for orphan conditions cannot be used to support Professor Hollis’ conclusion that his proposal to inhibit therapeutic class competition would increase development incentives. Instead, what the evidence on orphan drugs supports is a case against early generic entry.

Conclusions

Professor Hollis raises a number of interesting issues in his posting. However, his analysis of the effects on incentives for developing new drugs from creating higher regulatory hurdles for potential second and later entrants to a therapeutic class ignores or does not give due weight to such factors as the risks of failure, risk aversion, market expansion under rivalry, racing to horizontally differentiated niches, expected R&D costs, and expected delays in marketing. These factors argue strongly against his conclusion that inhibiting therapeutic class competition through creating regulatory barriers of the type he discusses will inevitably increase drug development incentives. In addition, although Professor Hollis looks to orphan drug policies and their impact on the development of drugs for orphan conditions to support his claim that innovation incentives will be increased if strong regulatory barriers to the approval of me-too drugs are erected, a careful examination of the purposes and nature of marketing exclusivity for orphan drugs shows that the orphan drug experience provides evidence for restricting early generic entry, not for restricting therapeutic class competition.

Professor Hollis’ proposed regulatory change is meant to apply to all drugs. While it is interesting to examine the implications of Professor Hollis’ proposal for the private sector development of drugs in general, it seems to me that such a change would do nothing to advance the cause of finding ways to encourage the discovery and development of drugs to treat neglected diseases. Here the essential problem for private developers is the absence of viable markets. Even a guarantee of a monopoly in research and a perpetual monopoly in the marketplace would not be sufficient to induce development if markets do not exist or are too weak.
References


