NOTE

IMPROVING POST-APPROVAL RISK SURVEILLANCE FOR DRUGS: ACTIVE POST-MARKET RISK IDENTIFICATION

Matthew Gordon[*]

Cite as: Matthew Gordon, *Improving Post-Approval Risk Surveillance for Drugs: Active Post-Market Risk Identification*
15 MICH. TELECOMM. TECH. L. REV. 297 (2008),
available at http://www.mttlr.org/volfifteen/gordon.pdf

Pre-approval clinical trials cannot possibly ensure that a drug will not have disastrous side effects once it arrives on the market. Post-approval drug safety data gathering was put in place to address this problem, but as implemented, it has not proven to be as effective as hoped. Congress recently overhauled the legislation regarding post-approval drug risk identification, and in doing so made a deliberate decision to put much of the burden of post-approval drug surveillance on the FDA through data mining. Further, the legislation gave the FDA the power to require post-approval clinical trials from drug makers only in limited circumstances. While this arrangement might seem wrong at first, the system, properly implemented, likely represents the most efficient option for risk identification at present. Still, to optimize the system, the FDA and HHS will have to cooperate to ensure that electronic health records are integrated into the data mining prospects. This active post-market risk identification system also has the potential to revolutionize other aspects of drug regulation, like off-label use, and the requirement for pre-approval clinical trials.

* J.D. expected May 2009, University of Michigan Law School. Ph.D. 2006, Molecular Biology, University of Utah. B.S. 2000, Biochemistry, Brigham Young University. msgordon@umich.edu.

I. INTRODUCTION

The Food and Drug Administration ("FDA"), created by the Food, Drug and Cosmetic Act of 1906 ("FDCA"),[1] has duties that range from ensuring that food products are not adulterated to approving the labeling of drugs shipped in interstate commerce.[2] One of the most famous, and arguably one of the most important roles of the FDA, is to ensure that drugs sold in the United States are safe and effective. The recent events surrounding the removal of Vioxx from the market[3] illuminate the need to consider additional ways to ensure the safety and effectiveness of drugs in the United States. This Note examines recent legislation which requires the FDA to execute a program of active post-market risk identification and proposes that mining data in electronic health records will optimize this system. This Note also compares the statutory active post-market risk identification system to a popular risk-surveillance alternative—requiring post-approval clinical trials for all drugs—and concludes that the active post-market risk identification system is superior. Finally, this Note identifies other aspects of drug regulation that may be affected by the recent legislation.

A. *FDA as Gatekeeper*

No drug may be introduced into interstate commerce unless it has been approved by the FDA.[4] The FDA's role as gatekeeper of the drug market was introduced in the 1962 amendments to the FDCA.[5] The drug approval process begins when a drug maker submits an Investigational New Drug ("IND") application after pre-clinical research identifies a promising drug candidate. With an FDA-approved IND application, the drug maker is authorized to begin clinical trials to test the safety and efficacy of the drug candidate.

The FDA requires three phases of clinical trials before a drug candidate can be approved for commercial use. In Phase I, the drug candidate

1. 21 U.S.C. §§ 301–99 (2006).
2. *See* History of the FDA, http://www.fda.gov/oc/history/historyoffda/default.htm (last visited Feb. 10, 2009).
3. *See* Gina Kolata, *MERCK and VIOXX: The Overview; A Widely Used Arthritis Drug Is Withdrawn*, N.Y. TIMES, Oct. 1, 2004, at A1.
4. Food, Drug and Cosmetic Act § 505(a), 21 U.S.C. § 355(a) (2008).
5. Act of October 10, 1962, Pub. L. No. 87-781, 76 Stat. 780 (1962).

is usually administered to healthy volunteers in order to determine appropriate dosage levels.[6] Studies in this phase will also gather data regarding pharmacokinetics, metabolism and side effects of taking the drug candidate.[7] Sometimes, preliminary data regarding effectiveness also come out of a Phase I trial.[8] If toxicity problems are identified in Phase I, the FDA may impose a clinical hold which prohibits the drug study from advancing to Phase II trials.[9]

Phase II clinical trials typically involve several hundred patients who have the condition that the drug candidate is intended to treat.[10] In most circumstances, Phase II trials are well-controlled and closely monitored.[11] In Phase II, safety data (including reports of adverse events) are also recorded. Typically, if Phase II data indicate that the drug candidate does improve the condition of patients, and no untoward safety concerns are raised, the drug candidate will be submitted to a final pre-approval trial in Phase III.[12]

A Phase III trial generally involves several hundred to several thousand patients who also have the condition the drug candidate is intended to treat.[13] This larger sample allows greater statistical power to determine whether the drug candidate alleviates symptoms and is safe to use at the effective doses. Phase III trials will sometimes compare the drug candidate to existing treatments, although the approval process does not currently require that a drug candidate be safer or more effective than existing treatments.[14]

If a company successfully shepherds its drug candidate through Phase I–III clinical trials, it can then submit a New Drug Application ("NDA") to the FDA for approval. The NDA must include the data that were generated in the clinical trials; the FDA examines this data to determine whether the drug candidate is safe and effective.[15] If the FDA

---

6. U.S. DEP'T OF HEALTH & HUMAN SERVS., THE CDER HANDBOOK 8 (1998), *available at* http://www.fda.gov/cder/handbook/.

7. 21 C.F.R. § 312.21 (2005).

8. *Id.*

9. U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 6, at 8.

10. 21 C.F.R. § 312.21.

11. U.S. DEP'T OF HEALTH & HUMAN SERVS., *supra* note 6, at 8.

12. *Id.* at 8–9.

13. 21 C.F.R. § 312.21(d) (2005).

14. *See* 21 U.S.C. § 355(d) (2008).

15. *See* 21 C.F.R. § 314.50–.53 (2008). In making this determination, the FDA balances the risks against the benefits. *See* Linda Bren, *Cancer Drugs: Weighing the Risks and Benefits*, 41 FDA CONSUMER MAG. (Jan–Feb. 2007), *available at* http://www.fda.gov/fdac/features/2007/107_cancer.html. In cases where the benefits are great or the therapeutic options are limited, greater risks will be tolerated. For example, many cancer drugs are also carcinogenic. Mikhail Blagosklonny, *Carcinogenesis, Cancer Therapy and Chemoprevention*, 12 CELL DEATH & DIFFERENTIATION 592, 592 (2005).

finds that the drug candidate is safe and effective, and other technical requirements are met, it will approve the NDA and the drug candidate can be legally marketed and sold in the United States.[16]

While the FDA approval process gives a drug candidate the imprimatur of hard science, there are obvious shortcomings with this risk assessment system. For instance, while clinical trials often involve hundreds or thousands of patients, these samples are still too small to statistically resolve some uncommon but disastrous side-effects.[17] Additionally, because Phase III trials generally last only one to four years, side effects that develop over a longer time period will not manifest themselves.[18] Furthermore, a growing chorus of researchers caution that women, minorities and other groups are grossly underrepresented in clinical trials.[19] This raises the possibility that safety and efficacy data generated in the trials may mislead doctors in treating the underrepresented groups.[20] Some also claim that the FDA approval process is flawed because of agency capture, or conflicts of interest with individuals involved in the approval process.[21] Finally, others argue that the pre-approval gatekeeping regime unduly delays the entry of important drugs onto the market.[22]

Various solutions have been proposed to fix these shortcomings, but one of the most prominent is the practice of post-market risk identification for drugs used by consumers.[23] While post-market surveillance of

16. *See* 21 U.S.C. § 355(c)(1), (d) (2008).

17. The most famous recent example is Vioxx, where the pre-approval data showed hints of cardiovascular side effects that were not statistically significant. It was not until a post-approval trial was conducted that the harmful side effects came to light. Kolata, *supra* note 3, at A1.

18. *See* MAYO CLINIC, CLINICAL TRIALS: A CHANCE TO TRY EVOLVING THERAPIES, http://www.mayoclinic.com/health/clinical-trials/DI00033 (last visited Feb. 11, 2009).

19. EDICT PROJECT, CHRONIC DISEASE PREVENTION AND CONTROL RESEARCH CENTER AT BAYLOR COLLEGE OF MEDICINE, MAJOR DEFICIENCIES IN THE DESIGN AND FUNDING OF CLINICAL TRIALS: A REPORT TO THE NATION IMPROVING ON HOW HUMAN STUDIES ARE CONDUCTED (2008), http://www.bcm.edu/edict/PDF/EDICT_Project_White_Paper.pdf.

20. It is unclear whether the underrepresented groups actually will respond differently. For example, race is generally considered a poor surrogate for genetic variability. *See* The HapMap Project, *Ethical Concerns*, http://www.hapmap.org/ethicalconcerns.html.en (last visited Feb. 11, 2009) ("[C]ommon ideas about race emerge largely from social and cultural interactions and are only loosely connected to biological ancestry.").

21. For example, drug advisory committees are composed of "experts" that recommend whether to approve a product based on data submitted with an NDA. USA Today reported that 54 percent of these experts had a financial stake in the outcome, including ownership of stock, research grants and consulting arrangements with the company submitting the NDA. Dennis Cauchon, *FDA Advisers Tied to Industry*, USA TODAY, Sept. 25, 2000, at A1.

22. *See* James L. Zelenay, Jr., *The Prescription Drug User Fee Act: Is a Faster Food and Drug Administration Always a Better Food and Drug Administration?*, 60 FOOD & DRUG L.J. 261 (2005).

23. *See* Marcia Crosse, Director Health Care, Drug Safety: Further Actions Needed to Improve FDA's Postmarket Decision-Making Process (May 9, 2007), *in* U.S. GOV'T AC-

drugs may patch some of the inherent shortcomings of the clinical trial system, prior to the enactment of the Food and Drug Administration Amendments Act of 2007 ("FDAAA"), it was a very inadequate patch.[24]

B. *Beyond Gatekeeping*

Before the passage of the FDAAA, the FDA's major post-market risk surveillance tools comprised: (1) mandatory Adverse Drug Experience ("ADE") reporting; (2) voluntary reporting of side effects by physicians; and (3) and post-approval clinical trials.

Mandatory ADE reports require drug makers to forward any information they receive regarding a drug's harmful side effects to the FDA.[25] If the adverse effect is "serious and unexpected," the drug manufacturer must report this information within 15 days and promptly investigate the event.[26] Otherwise, drug manufacturers submit quarterly ADE reports for the first three years post-approval, followed by yearly reports thereafter.[27]

While helpful, this system is criticized as not being clear and effective, and not adequately establishing the actual extent of drug safety problems.[28] For example, drug companies have a very strong financial incentive to delay the release of ADE data, avoid classifying adverse events as "serious and unexpected," or not submit data at all.[29] The FDA has been criticized for failing to bring actions against companies that do not submit ADE data in a timely manner.[30]

The second tool used for post-approval drug surveillance—independent reporting by practitioners—is also subject to criticism. First, many doctors fail to report side effects. For example, doctors may not recognize that a common ailment (e.g., a heart attack) is actually linked to the use of a certain drug, or they may simply lack the infrastructure or incentives to consistently report adverse events. Unfortunately, it is estimated that medical practitioners report only about 10 percent of harmful

---

COUNTABILITY OFFICE, GAO-07-856T, 2007, at 3–4, *available at* http://www.gao.gov/new.items/d07856t.pdf.

24. It is possible that the current post-market risk-identification regime is the most efficient system, despite the fact that many harmful drugs are not pulled from the market in a timely manner. Still, Congress' amendments to the system are an indication of the popular opinion that post-market risk identification could function better. *See also id*. at 2.

25. Food, Drug and Cosmetic Act § 505(k)(1), 21 U.S.C. § 355(k)(1) (2008); 21 C.F.R. § 314.80(a) (2008).

26. 21 C.F.R. § 314.80(c) (2008).

27. *Id.*

28. Crosse, *supra* note 23, at 8.

29. Failing to classify an adverse event as "serious and unexpected" would allow the drug company to delay reporting of the event.

30. *See* Marc Kaufman, *Reports to FDA Were Late, Watchdog Says*, WASH. POST, Aug. 4, 2004, at E2.

drug side effects,[31] which reduces the likelihood that the FDA will recognize the correlation between the drug and the harmful effect. Indeed, even when reports are made, the FDA may not act on them unless it receives similar reports from many practitioners.[32]

Post-approval, or Phase IV, clinical trials represent the third and most analytically powerful pre-FDAAA tool for identifying drug risks. Phase IV trials are generally longer than pre-approval trials, and can include a much broader range of individuals,[33] thus alleviating some of the shortcomings of pre-approval trials. Yet, these trials were not a viable source of post-approval risk data because the FDA lacked the power to compel drug manufacturers to complete them.[34] The FDA could extract promises from drug companies to complete Phase IV studies, but once approval was granted, the financial incentive to complete the trials in a timely manner disappeared. As a result, many Phase IV trials are still pending.[35]

In conclusion, pre-FDAAA post-approval risk-assessment tools were inadequate. In the FDAAA, Congress addressed this problem by requiring the FDA to implement a risk identification system based on data-mining, and allowing the FDA to require Phase IV clinical trials when other methods of risk assessment are deemed inadequate. The balance of this paper will: (1) describe the new regime of post-market risk identification; (2) explore how it might be optimally implemented; and (3) compare it economically to a popular risk-assessment alternative.

II. THE NEW REGIME OF POST-MARKET RISK REGULATION

On September 27, 2007, the FDAAA was signed into law.[36] It constitutes one of the largest revisions of the FDCA in several decades and deals with many issues. In particular, some of the most interesting

---

31. *See* Marc Kaufman, *Drug Safety Panel is Criticized*, WASH. POST, June 8, 2005, at A5.

32. *See* Geeta Anand, *Jaw Ailment Shows Industry Moves Slowly on Drug Warnings*, WALL ST. J., Dec. 8, 2004, at B1 (describing the experience of a doctor who tried for two years to convince the FDA that a cancer drug was causing osteonecrosis in patients).

33. Lisa R. Johnson-Pratt, *Phase IV Drug Development: Post-Marketing Studies*, *in* PRINCIPLES AND PRACTICE OF PHARMACEUTICAL MEDICINE 119 (Andrew J. Fletcher et al. eds., 2d ed. 2007).

34. The FDA now has the authority to require clinical trials, but only under the limited circumstances discussed below. Food, Drug and Cosmetic Act § 505(o)(3)(D), 21 U.S.C. § 355(o)(3)(D) (2008).

35. *See 65% of Promised Drug Studies Pending*, WASH. POST, Mar. 4, 2006, at A4.

36. Food and Drug Administration Amendments Act of 2007, Pub. L. No. 110-85, 121 Stat. 823 (2007).

changes to the FDCA address the problems of post-market risk identification.

A. *Active Post-Market Risk Identification*

The changes to post-approval risk identification are found in Title IX of the FDAAA, "Enhanced Authorities Regarding Post-market Safety of Drugs." In particular, Sections 901 and 905 make major structural changes to the actual process of post-market risk surveillance. Section 901, which amends Section 505(o) of the FDCA, authorizes the FDA to require that drug makers perform post-market, or Phase IV clinical trials, but only under limited circumstances. The FDA may only require Phase IV clinical trials when the Secretary of the Department of Health and Human Services ("HHS")[37] finds that: (1) post-market risk identification provisions in FDCA Sections 505(k)(1) and (3) are inadequate; and (2) a less rigorous "study" is likewise inadequate.[38] This will not please those who advocate mandatory Phase IV trials of all approved drugs.[39]

Sections 505(k)(1) and (3) list the preferred post-market risk identification systems. Section 505(k)(1) sets out the requirement, mentioned previously, that a drug manufacturer keep a record of and report the negative clinical outcomes from post-approval drug use.[40] On the other hand, Section 505(k)(3) was added by Section 905 of the FDAAA, and represents a new approach to post-market risk identification. It requires the FDA to "develop methods to obtain access to disparate data sources,"[41] and also develop and implement a "postmarket risk identification and analysis system" based on those methods.[42] The methods must be developed by September 2009, and the system must be implemented no later than September 2010.[43] The overall goal is to "link and analyze safety data from multiple sources,"[44] including "claims data, patient survey data, standardized analytic files that allow for the pooling and analysis of data from disparate data environments, and any other data deemed appropriate by the Secretary."[45]

Section 905 of the FDAAA also adds Section 505(k)(4) to the FDCA, directing the FDA to collaborate with public and private entities

37. The Secretary does not actually make this decision, but will delegate it to the FDA.
38. Food, Drug and Cosmetic Act § 505(o)(3)(D), 21 U.S.C. § 355(o)(3)(D) (2007).
39. Mark Greener, *Drug Safety on Trial*, 6 EMBO REPORTS 202, 202 (2005).
40. Food, Drug and Cosmetic Act § 505(k)(1); 21 C.F.R. § 314.80(a).
41. Food, Drug and Cosmetic Act § 505(k)(3)(B)–(C).
42. *Id.*
43. *Id.*
44. *Id.* § 505 (k)(3)(B).
45. *Id.* § 505(k)(3)(A).

to perform advanced query analyses with the data system.[46] The FDA can also contract for the same services.[47] Congress has instructed the FDA to include in these analyses data from at least 25 million patients by 2010 and 100 million patients by 2012[48]—nearly eighteen-thousand times as many individuals as the mean total enrollment in pre-clinical trials.[49]

The appeal of this approach is immense: if sufficient access to quality data is secured, this system becomes analogous to a continuous, massive post-approval "clinical trial" with the potential for ongoing feedback into the risks of the drug in practice. Additionally, the FDA will be guiding the research, so gamesmanship by the regulated entities will not potentially muddle data. Further, the potential scope of the data analysis is so large that statistical analyses will be able to identify low probability risks.

While the system promises many benefits, it will also incur costs. The FDA must develop complex methods and systems for data gathering and analysis. Because of the enormity of the data set, and the overwhelming number of potential drug safety queries, these data-mining activities could be a bottomless pit of resource consumption. To avoid overspending, Congress requires that drug safety questions be addressed in a prioritized order established biannually through consultation with the Drug Safety and Risk Management Advisory Committee or its successor.[50] Still, this does not necessarily limit the costs; it only establishes which risks are most likely to be addressed.

The true question, then, is whether the new regime maximizes the potential net benefits[51] available through post-approval risk identification systems. In response, this paper will first explore how the active post-market risk identification system established by Congress in Section 905 of the FDAAA can be optimized. This "optimal" system will then be compared to the other main post-approval risk identification system: mandatory Phase IV clinical trials for all drugs.

B. *Optimizing the Congressional Mandate*

Congress gave very few specific instructions to the FDA regarding (1) which data sources to mine; (2) how to ensure ongoing access to up-

46. *Id.* § 505(k)(4)(A). For example, the FDA might wonder whether a recently approved drug causes an increased incidence of a certain condition in users. It could query the health data to see if there is a correlation between the two.

47. *Id.* §§ 505(k)(3)(C)(v), 505(k)(4)(D), 505(k)(4)(F), 505(k)(4)(G).

48. *Id.* § 505(k)(3)(B)(ii).

49. Joseph A. DiMasi et al., *The Price of Innovation: New Estimates of Drug Development Costs*, 22 J. HEALTH ECON. 151, 177 n.41 (2003).

50. Food, Drug and Cosmetic Act § 505(k)(4)(C).

51. The total benefits of the system minus its costs of implementation, and any other incidental costs.

dated drug safety information; and (3) which data-mining techniques to use in drug safety queries. The third issue is a concern for database and information processing professionals. For the purpose of this paper, I assume that the FDA will choose the appropriate data-mining and analysis techniques for each risk question that is presented. This leaves the task of optimizing the first two issues.

1. Data Sources

The FDAAA lists two potential sources of data for the active post-market risk identification system: (1) "federal health-related electronic data (such as data from the Medicare program and the health systems of the Department of Veterans Affairs);" and (2) "private sector health-related electronic data (such as pharmaceutical purchase data and health insurance claims data)."[52] Beyond this, the statute merely directs the FDA to choose "other data as the Secretary deems necessary to create a robust system . . . ."[53] Both potential sources specifically mentioned comprise mostly health insurance account data or analogous government files.

Health insurance account data contains records of which drugs were prescribed to an individual for a particular condition. An insurance account file will also include claims for diagnoses of any conditions that arise after the drug is prescribed to the patient. Thus, if the FDA receives several ADEs reporting an adverse reaction associated with a certain drug, the FDA could then search insurance claim data to determine in general whether individuals who are prescribed the drug are more likely to be diagnosed with the putative adverse side effect.[54]

Still, there are drawbacks to mining health insurance claim records for safety analysis. Most importantly, insurance claim data do not include positive outcomes from use of a drug.[55] This is unfortunate because the FDA typically assesses the safety and efficacy of a drug by balancing its harms against its benefits. Thus, even if a drug risk is identified, the FDA may still not have enough information to decide whether to allow the drug

52. Food, Drug and Cosmetic Act § 505(k)(3)(C)(i)(III).

53. *Id.*

54. Of course, limits would have to be put on the search. For example, the search would likely require that the drug in question was prescribed recently to improve the likelihood of seeing causal connections. These limits are specific to the data-mining and statistical analysis process, and as stated before, I assume that the FDA will strive to optimize its search algorithms.

55. For example, if I am treated with a drug for a condition, the insurance reflects this, but my insurance provider will likely not be notified if my condition improves with treatment. The only evidence that I was healed would be a cessation of treatment.

to stay on the market.[56] For example, imagine a drug that is prescribed to treat childhood leukemia that was originally approved by the FDA for treatment of breast cancer in adults. No children were enrolled in the pre-approval clinical trials. However, due to very encouraging results (a few cases of miraculous recovery) from off-label use in some children, the drug became widely prescribed in children. One doctor reports a fatal aortic aneurism in a pediatric patient after treatment with the drug. A data-mining query confirms an elevated risk of fatal cardiovascular events in children with use of the drug. Here, the FDA is faced with a confirmed risk, but only anecdotal information regarding benefits, so a traditional risk-benefit analysis will be stymied.

This might seem to imply that data-mining will not be very effective in general for post-market risk identification; however, this is not the case. The problems with insurance claim data can be avoided. HHS is currently following a worldwide trend of pushing the incorporation of Electronic Health Records ("EHRs") into American healthcare.[57] President Bush set a "goal for most Americans to have access to secure electronic health records by 2014."[58] EHRs contain all of the information regarding a patient under a physician's care. While they include all of the information found in a typical insurance claim record, they are much richer. EHRs include the results from tests performed to confirm diagnoses, as well as positive outcomes from treatment.[59]

If EHRs could be incorporated as a data source in post-market risk surveillance,[60] data-miners would have access to the same types of information as obtained in typical clinical trials. Thus, data-miners would be able to ascertain to some degree the benefits as well as the risks of a

---

56. The FDA might alternatively require labeling changes or other solutions if it could assess the risk-benefit ratio, but it is prohibited from doing so.

57. The present justification for the use of EHRs is to correct problems with current medical record-keeping techniques. Generally, an individual will have a health record on file with multiple providers, but EHRs are designed to centralize and standardize medical record keeping. EHRs will be portable, and should reduce (sometimes fatal) errors in medical care while increasing transparency in the industry generally. U.S. Dept. of Health and Human Servs., Nationwide Health Information Network (NHIN): Background, http://www.hhs.gov/healthit/healthnetwork/background/ (last visited Feb. 14, 2009).

58. U.S. Dept. of Health and Human Servs., Health Information Technology, American Health Information Community, http://www.hhs.gov/healthit/community/background/ (last visited Feb. 14, 2009). The system imagines a nationwide health information network, which will "provide a secure, nationwide, interoperable health information infrastructure that will connect providers, consumers, and others involved in supporting health and healthcare." U.S. Dept. of Health and Human Servs., *supra* note 57.

59. *See* Richard Bordowitz, *Electronic Health Records: A Primer*, 39 LABMEDICINE 301, 301 (2008).

60. EHRs would need to be stripped of all identifying information, as insurance claim data would also have to be. Indeed, most risk-data sources raise HIPAA privacy rule issues, but these are not insurmountable if the right methods are built in from the beginning.

drug in a large segment of all patients taking the drug. HHS is attempting to guide the uptake of EHRs so that they will be machine readable and based on compatible standards.[61] This common information architecture will allow easier access to a massive number of records for data-mining.

2. Ongoing and Continuous Access

A successful active post-market risk identification system will also optimize the availability of ongoing and continuous access to data. A potential problem with insurance claim data as a source of data-mining is the plethora of insurance providers. An individual may change providers for many reasons: to save money, because coverage with the current provider is not available after moving to another region, or merely because his employer decides to change plans. All of these scenarios will destroy the information continuity that is required to establish and maintain a robust active post-market risk identification system. While this discontinuity in insurance claim records is likely not a fatal flaw to the system, it will undoubtedly hamper the ability to find real risks by introducing false negatives.[62]

Fortunately, this difficulty can be ameliorated by introducing EHRs as a data source. EHRs will be implemented in a way to ensure that the record remains with a patient instead of a provider. Thus, even if an individual moves across the country, the EHR containing all of the individual's medical history will be available to his new doctors. This ongoing and continuous record of data will also allow FDA data-miners to query whether particular side effects are more common in long-term users of certain drugs.

In sum, optimizing the active post-market risk identification system will require that HHS and the FDA cooperate to ensure that EHRs become available as a data source. Since both programs are set to be accomplished over the same period,[63] the timing could not be better for cooperation between the two projects.

---

61. *See* U.S. Dept. of Health and Human Servs., *supra* note 57. *See also* Exec. Order 13,410, 3 C.F.R. § 240 (2007), *proposed amendment in* 42 U.S.C.A. § 300 (2008).

62. For example, a person may begin taking a drug and charge the drug to his first insurance provider. After the drug cycle is completed, the individual changes insurance providers and experiences a side effect of the drug that is billed to the second insurance carrier. There will be no linkage in the data between the drug prescription and the adverse event.

63. The post-market risk identification system is slated to include 100 million patients by 2012, Food, Drug and Cosmetic Act § 505(k)(3)(B)(ii), and EHRs are to be available to most Americans by 2014. *See* U.S. Dept. of Health and Human Servs., *supra* note 57.

C. *"Data-Mining First" vs. "Required Phase IV Trials"*

Despite passage of the FDAAA, some still argue that all drugs should be provisionally approved subject to a requirement that Phase IV clinical trials be performed in a timely manner.[64] Thus, it will be helpful to compare a required Phase IV ("RPIV") regime to the one laid out in the FDAAA, which requires data-mining first ("DMF"). Recall that the latter will still allow the FDA to require Phase IV trials, but only if it is determined that the data-mining system is found to be inadequate.[65]

1. Financing the Systems

Although DMF and RPIV appear at first to be financed in very different ways, with either system, the American public foots the bill. As currently legislated by Congress, DMF will be funded by the FDA—meaning tax revenue will pay for the increase in drug safety. On the other hand, RPIV expenses incurred by drug companies would be passed on to drug consumers in the form of higher drug prices. The higher drug prices would be spread over the American public through higher insurance premiums. Assuming most Americans carry health insurance and pay taxes, either system will pass the costs on to the public.[66]

Since the costs will be distributed over the American public regardless of the mechanism of financing, the total costs of each system will be compared next.[67] A key assumption for the purpose of this comparison is that the director of the FDA will correctly recognize when data-mining is insufficient to identify public health risks, and a Phase IV clinical trial is required. Under that assumption, the public health outcomes in DMF and RPIV would be the same, and public health costs and benefits do not need to be factored into the comparison.

By factoring out hard-to-quantify public health costs and benefits it is possible to express the cost differential between RPIV and DMF

64. *See, e.g.*, Greener, *supra* note 39, at 202.

65. As mentioned previously, *supra* note 24, there is a technical possibility that the system in place before passage of the FDAAA is more efficient than either option discussed in this section. Besides being politically unpalatable, the pre-FDAAA system's relative cheapness is likely outweighed by its failure to identify many harmful drugs. A formal comparison of DMF and RPIV to the previous system is outside the scope of this paper.

66. This is overly simplistic, as health insurance and taxes do not distribute costs identically. Still, while any individual is unlikely to bear the same fraction of the costs under the system, for the purposes of this economic analysis, the only consideration is which regime has a greater aggregate cost to society.

67. It may be more politically tenable to have the drug companies "pay their way," thus favoring the RPIV model where the drug companies appear to be bearing the costs. Indeed, DMF might be further optimized by amending the statute to require any drug company with a drug on the market to pay a percentage of their gross revenue on the drug to fund the active post-market risk identification system. This would remove the system's funding from the vagaries of budget appropriations.

mathematically. Under DMF, Phase IV clinical trials will still be performed, but only when data-mining fails. Given that the variable T represents the total cost of RPIV (i.e., the cost of performing a Phase IV trial on every approved drug), and the variable t represents the total cost of the clinical trial component of DMF,[68] then the difference in clinical trial costs between these systems (expressed as ΔC_{trials}) will be:

$$(1) \qquad \Delta C_{trials} = T - t$$

This can be rewritten as:

$$(2) \qquad t = T - \Delta C_{trials}$$

Let M represent the cost of the data-mining component of DMF, so that the total cost of DMF can be expressed as $M + t$. Since the total cost of RPIV is T, the difference between the total cost of RPIV and the total cost of DMF (expressed as ΔC_{total}) will be:

$$(3) \qquad \Delta C_{total} = T - (M + t)$$

Now, substituting equation (2) for t in equation (3):

$$(4) \qquad \Delta C_{total} = T - (M + T - \Delta C_{trials})$$

This simplifies to:

$$(5) \qquad \Delta C_{total} = \Delta C_{trials} - M$$

Under the assumptions stated above, if ΔC_{total} is positive, DMF is cheaper to implement, and vice versa. ΔC_{total} is only positive (and DMF the cheaper option) when $M < \Delta C_{trials}$.[69]

Under a further set of assumptions, it is possible to perform a concrete analysis determining whether DMF is less expensive than RPIV. First, assume that the average cost of a Phase IV clinical trial is $86 million,[70] or about $0.75 per household per trial.[71] Second, the analysis will

68. T will always be greater than t unless the data-mining component of DMF is unable to identify any post-approval drug risks and the FDA is forced to require Phase IV clinical trials for every drug anyway. In such a case, $T = t$.

69. Intuitively this makes sense: if the cost of data-mining (M) does not consume the savings of having fewer clinical trials under DMF ($\Delta Ctrials$), then DMF is cheaper than RPIV.

70. *See* DiMasi et al., *supra* note 49, at 162 (concluding that the average cost of a Phase III clinical trial was $86 million, inflation-adjusted to the dollar value in the year 2000). Given the scarcity of data regarding the actual cost of a phase IV clinical trial, the figure for a Phase III trial will be used.

be restricted to new molecular entity ("NME") approvals only, not reformulations of previously approved drugs.[72] Third, assume that 30 NMEs will be approved per year.[73] Thus, RPIV would cost roughly $2,580,000,000, or $22 per household.[74] This represents the total cost of RPIV (T in formulas 1–4) under these assumptions.

With regards to DMF, it can be assumed that half of all drugs approved by the FDA will have an adverse side effect.[75] Additionally, it can be assumed that data-mining will, at a minimum, flag all drugs with adverse side effects as potentially harmful, but that the rate of false positives is moderate (e.g., 20 percent of all approved drugs are inappropriately flagged as potentially harmful).[76] Further, it can be assumed that data-mining is unable to determine the risk/benefit ratio for drugs flagged as suspicious.[77] Therefore, a Phase IV clinical trial will be required whenever a risk is identified under DMF. Under these assumptions, ΔC_{trials} becomes the cost of performing clinical trials on 30 percent of all approved NMEs,[78] or $774 million per year.[79]

Recall that under equation (5), it was concluded that DMF is less costly than RPIV when $M < \Delta C_{trials}$. Under the assumptions above, DMF

---

71. This is to provide an idea of the individual cost burden of a clinical trial. There were 114,384,000 households in the United States in 2006. U.S. Census Bureau, America's Families and Living Arrangements: 2006, at tbl.AVG1, http://www.census.gov/population/www/socdemo/hh-fam/cps2006.html (last visited Feb. 14, 2009). Thus, if a phase IV trial costs $86 million and there are 114,384,000 households, the cost per household is $86,000,000 divided by 114,384,000, or $0.75.

72. NMEs will benefit most from an active post-market risk identification system as there has been little clinical experience with them. On the other hand, an extended release form for a well-known drug is less likely to have unknown harmful side effects.

73. *See* TUFTS CENTER FOR THE STUDY OF DRUG DEVELOPMENT, OUTLOOK 2006 at 1 (2006), *available at* http://csdd.tufts.edu/InfoServices/OutlookPDFs/Outlook2006.pdf (about 30 NMEs were approved per year on average in the period from 1999–2004).

74. Remember, it was assumed that the costs would be passed on to consumers either through higher taxes or higher insurance premiums. The per household number is obtained by dividing the total cost, $2,580,000,000, by 114,384,000 total households.

75. This assumption is based on a study which concluded that over half of all drugs approved between 1976 and 1985 had serious side effects which were discovered only after approval. U.S. GEN. ACCOUNTING OFFICE, FDA DRUG REVIEW: POST APPROVAL RISKS 1976–85 at 3 (1990).

76. This is a reasonable estimate because higher sensitivity in any detection system generally leads to higher rates of false positives.

77. This is a conservative estimate of the capability of the post-market risk identification system. Hopefully, the results will be better. As explained above, by using EHRs the FDA may be able to perform enough risk/benefit balancing to avoid a clinical trial in some cases. Making this assumption allows for the determination of the minimal cost savings of DMF relative to RPIV. A less conservative assumption will yield even greater savings for DMF.

78. The harmful drugs (50 percent of the total), plus the false positives (20 percent of total) result in 70 percent of all drugs needing clinical trials under DMF. Thus, the costs of 30 percent of trials under RPIV are saved under DMF.

79. If T = $2,580,000,000 per year, *supra* note 74, and $\Delta Ctrials = 0.3T$, then $\Delta Ctrials$ = $774,000,000 per year.

is the more economically efficient system of post-approval drug risk surveillance if the yearly cost of its data-mining component (*M*) is less than $774 million.

Unfortunately, predicting the costs of implementing FDAAA's active post-market risk identification system is difficult due to a dearth of pricing information regarding the implementation of data-mining schemes. It is telling, however, that the 2008 FDA budget for modernizing drug safety is set at only $11.2 *million*.[80] This is a roughly 70-fold difference. Indeed, due to the heavy use of computers in the DMF system, it is likely that many operations could be automated. Thus, after initial costs of implementation are invested, it is highly unlikely that the annual operating budget could surpass hundreds of millions of dollars. In this light, it appears that DMF is the cheaper alternative.

2. Structural Safeguards

Safeguards built into the mechanics of these systems can ensure that strategic gaming and human error do not compromise the systems' goal of identifying post-approval risks. By comparing the safeguards incorporated in DMF and RPIV, the relative utility of the two systems can be determined.

DMF will be administered by the FDA, a government agency, so it is less likely that there will be financial pressure for the risk-surveying entity to "cheat." RPIV, on the other hand, is administered by the drug companies which have a financial stake in generating data consistent with safety. While it is true that RPIV data would be reviewed by professionals at the FDA, drug companies have shown a particular adeptness in discovering ways to maximize profits under the FDCA's regulatory scheme.[81] It can be assumed that they are equally imaginative in performing clinical trials. This is not meant to imply that drug companies are corrupt—only that they will generate the "best possible" safety data. While drug companies have some incentive to generate honest data to reduce potential tort liability for harmful drugs which remain on the market, since the federal preemption question regarding tort liability for prescription drugs is unsettled,[82] the strength of a drug manufacturer's

80. U.S. Food and Drug Administration, Summary of FDA's FY 2008 Budget, http://www.fda.gov/oc/oms/ofm/budget/2008/summary.html (last visited Feb. 14, 2009). In fact, the FDA would have to spend 37 percent of its entire budget request for 2008 of $2.1 billion on DMF to match the estimated cost of RPIV.

81. For example, patent evergreening. *See* Michael Enzo Furrow, *Pharmaceutical Patent Life-Cycle Management After* KSR v. Teleflex, 63 FOOD & DRUG L.J. 275, 276–77 (2008).

82. *See, e.g.*, Wyeth v. Levine, No. 06-1249, currently pending before the Supreme Court. Of course, the eventual decision in *Wyeth* may disambiguate preemption law to a great extent.

tort-based incentive is unclear. On balance, while the FDA lacks incentives to "cheat" while administering DMF, drug makers have at least a slight incentive to "massage" clinical trial data they generate under RPIV.

Although the FDA lacks incentives to "cheat," under DMF it also lacks compelling incentives to be as efficient as possible.[83] Under RPIV, on the other hand, drug companies face strong pressure to complete Phase IV trials in a cost-effective manner.

Perhaps the most important structural benefit of DMF is its scope. Data-mining is intended to eventually cover 100 million patients, which would represent about one-third of all Americans. Because it will include so many individuals, it is likely that data from every kind of user will be available and greater statistical resolution can be achieved. This eliminates many shortcomings of clinical trials addressed in Part I.A. Indeed, one reason clinical trial data are generated under well-observed and standardized procedures is to ensure that the data are "clean" and easier to interpret, because the results must be extrapolated from a small sample to the whole population. With the data-mining scheme legislated in the FDAAA, such a large sample will be queried that very little extrapolation will be required.

In conclusion, while drug companies are more likely to efficiently implement a post-market risk identification system, they also have at least a slight incentive to "massage" clinical trial data they would generate under RPIV. On the other hand, while the FDA would likely experience less pressure to either cheat or implement the system quickly and efficiently, the large size of the data-mining sample tips the balance of structural safeguards in favor of DMF.

III. CONCLUSION

In conclusion, if the DMF is implemented properly, it should be a better policy choice than RPIV for post-approval risk identification. First, it is less expensive to implement. This conclusion, however, is based on many assumptions, the most important of which is that the FDA, on its limited budget, can achieve substantially the same public health benefits using either system. While this assumption is difficult to prove, it is not unreasonable. Budget limitations would be less disruptive to DMF if the FDA was allowed to charge fees to drug companies as payment for the risk monitoring services they are providing. While costs will likely be passed on to taxpayers/insureds anyway, it is likely that

83. The only pressure would be political pressure.

fees would be a politically more tenable source of funding than taxes.[84] Second, while both systems have potential structural problems, the sheer scope of the data-mining promises to ensure that the conclusions drawn from DMF are more "honest."

Besides the allure of an efficient and better method of post-approval risk identification, the data-mining component of DMF has the potential to dramatically affect other aspects of drug regulation. For example, if, as proposed in this paper, EHRs are included as a data source and the tasks of data gathering and analysis can be automated, it might be possible to assess a drug's reactions in the public in "real time." While this capability may not be economical using current technology, as internet connection speeds and computing capacity increase, a system of truly active risk assessment might be possible. In fact, if the Health Information Technology Group at HHS and the FDA cooperate on the implementation of EHRs, it will be possible to organize EHRs so that automated scripts can search for both risks and positive outcomes from new drugs. Predefined scripts could automatically screen many or all approved drugs for a set of determined side effects.

This ability to gather risk/benefit data in real time would likely transform the process of drug regulation—and raise many interesting questions as well. First, it would allow for validation of off-label uses that appear promising in the medical literature. Still, would this validating data be made available to drug companies for marketing purposes? In particular, if the system is publicly funded, this could be seen as a huge subsidy to the drug industry.

Second, could the data-mining system, if properly validated, further replace some of the clunky pre-market clinical trial process? If reliable data regarding the safety and effectiveness of a drug can be obtained while the drug is on the market, perhaps only limited safety trials would need to be performed before a drug is approved. This would allow terminally ill patients access to promising new drugs as early in the drug life-cycle as possible. Admittedly, if this avenue is eventually explored, regulators will need to strike an appropriate balance to ensure that drugs allowed on the market will have at least a baseline level of safety.

Third, might a data-mining system incorporating EHRs change the types of drugs companies pursue for approval? Pharmacogenomics[85] promises to take us into an era of personalized medicine. This transition will be difficult, however, because the current system arguably incentiv-

84. *See supra* note 67.

85. Pharmacogenomics is a fledgling field that attempts to use genetic indicators in patients to estimate a patient's response to a particular drug.

izes drug companies to pursue only blockbuster drugs.[86] If drugs can enter the market with fewer pre-market safety data requirements, drug makers may be more willing to invest in smaller market genetically-tailored drugs. Furthermore, as genotyping becomes more common, EHRs will likely include a patient's genetic information and FDA data-miners could then look for correlations between certain genotypes and outcomes from drug use.

Whatever the eventual fate of data-mining post-market drug surveillance, this legislative step will not only improve drug safety, but take us to the brink of some very intriguing drug regulation measures.

86. The argument posits that the very high cost of the approval process will only be undertaken when the drug is likely to generate huge revenues.