

NBER WORKING PAPER SERIES

BLACK BOX WARNINGS AND DRUG SAFETY:
EXAMINING THE DETERMINANTS AND TIMING OF FDA WARNING LABELS

Allan Begosh
John Goldsmith
Ed Hass
Randall W. Lutter
Clark Nardinelli
John A. Vernon

Working Paper 12803
<http://www.nber.org/papers/w12803>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2006

The views expressed in this paper are those of the authors and do not necessarily reflect those of the institutions with which the authors are affiliated. Vernon is also a professor at the University of Connecticut and a Faculty Research Fellow with the National Bureau of Economic Research (NBER).

© 2006 by Allan Begosh, John Goldsmith, Ed Hass, Randall W. Lutter, Clark Nardinelli, and John A. Vernon. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

ABSTRACT

Comparing the safety of prescription drugs over time is difficult due to the paucity of reliable quantitative measures of drug safety. Both the academic literature and popular press have focused on drug withdrawals as a proxy for breakdowns in the drug safety system. This metric, however, is problematic because withdrawals are rare events, and they may be influenced by factors beyond a drug's safety profile. In the current paper, we propose a new measure: the incidence and timing of Black Box Warnings (BBWs). BBWs are warnings placed on prescription drug labels when a drug is determined to carry a significant risk of a serious or life-threatening adverse event. Using a unique data set, one that includes all new molecular entities (NMEs) submitted to the FDA between May 1981 and February 2006, and subsequently approved and marketed, we analyze the timing and incidence of BBWs. Our analyses also use data on several drug characteristics likely to affect the probability a new drug will receive a BBW. We draw several conclusions from our analyses. For example, drugs receiving priority FDA review are more likely to have BBWs at the time of approval than NMEs receiving standard review. We also find that early prescription volume and orphan drug status are associated with an increased likelihood of receiving a BBW. We do not, however, find a significant difference in the rate of BBWs across time cohorts. A comparison of NMEs approved before and after the 1992 Prescription Drug User Fee Act (PDUFA), which authorized the payment of user fees from drug manufacturers to the FDA in an effort to expedite new drug application (NDAs) review times, did not reveal a statistically significant difference in the rate of BBWs. Critics of PDUFA maintain that reduced FDA-approval times under PDUFA have compromised drug safety. We do not find empirical support for this contention.

Allan Begosh
Suite 15-36
Office of the Commissioner
Food and Drug Administration
5600 Fishers Lane
Rockville, MD 20857
allan.begosh@FDA.gov

Randall W. Lutter
Suite 15-36
Office of the Commissioner
Food and Drug Administration
5600 Fishers Lane
Rockville, MD 20857
Rlutter@oc.fda.gov

John Goldsmith
Suite 15-36
Office of the Commissioner
Food and Drug Administration
5600 Fishers Lane
Rockville, MD 20857
jcgoldsm@erols.com

Clark Nardinelli
Suite 15-36
Office of the Commissioner
Food and Drug Administration
5600 Fishers Lane
Rockville, Maryland 20857
Clark.Nardinelli@FDA.HHS.GOV

Ed Hass
Suite 15-36
Office of the Commissioner
U.S. Food and Drug Administration
5600 Fishers Lane
Ed.Hass@FDA.HHS.GOV

John A. Vernon
University of Connecticut
Department of Finance
2100 Hillside Road
Storrs, CT 06269
Office of the Commissioner, US FDA
and NBER
jvernon@business.uc

I. INTRODUCTION

For a new drug to gain FDA approval in the United States it must be demonstrated safe and effective when used as directed. While there is an abundance of clinical data on each new drug approved, there are no generally accepted and easily quantifiable measures of *overall* drug safety to track trends in drug safety over time or across therapeutic classes. Research efforts directed towards the identification of aggregate drug safety measures could not be timelier. According to a recent report by the Institute of Medicine (2006), there is a growing perception that we are in the midst of a drug safety crisis in the U.S., and that significant steps are required to improve the current system. The IOM report implies that a reliable, aggregate measure of drug safety in the U.S. could be very helpful in monitoring the performance of the current drug safety system.

In the absence of good proxies for system-wide drug safety trends, academic researchers and the popular press have been forced to focus primarily on drug withdrawals, which may be unreliable measures due to their infrequency. Safety withdrawals are drastic actions; they capture only the most serious lapses in drug safety. Of 516 new molecular entities (NMEs) approved by the FDA between 1981 and 2006, for example, only 16 were withdrawn for safety reasons.¹ Furthermore, a decision to withdraw a drug may be influenced by non-safety factors, such as the availability of

¹ A New Molecular Entity or NME is a drug that contains an active substance that has never before been approved for marketing, in any form, in the United States.

alternative treatments and the potential impact on patient welfare.² As no prescription drug is perfectly safe (which is why they are only available through a prescription), a better metric of overall safety should reflect actions to promote safety for a larger sample of drugs, including drugs with safety concerns that remain on the market.

The small number of safety withdrawals also limits the statistical utility of using withdrawals as a metric for drug safety. Meaningful statistical comparisons of withdrawal rates across cohorts are virtually impossible. As previously mentioned, of 516 NMEs approved for marketing by the FDA between May 1981 and February 2006 (the time period covered by our sample), there have only been 16 safety withdrawals, or about 3.1% of all NMEs. The corresponding 95% confidence interval associated with this observation is bounded on the interval from 1.6% to 4.6%, or between eight and 24 safety withdrawals.³ While the FDA, the general public, and other stakeholders would be gravely concerned with a system-wide failure or problem that increased the safety risk of drugs by 40%, an increase of this magnitude in safety withdrawals in the current sample would not be statistically significant at the 0.05-level. Further complicating potential measures of drug safety over time is the vintage of the drug, or drug cohort, post-approval. Older drugs have had more time to experience safety withdrawals; a fact that we will explicitly control for in our empirical analyses.

Safety withdrawals are not the only possible action that can be taken when safety concerns arise with a drug. A sponsor may elect to add, or the FDA may require them to add, a warning on the product labeling. For the most serious safety problems, which may

² One example of this would be the decision by the FDA to withdraw mibefradil and bromfenac but to keep troglitazone. See Fung, Man, "Evaluation of the Characteristics of Safety Withdrawal of Prescription Drugs from Worldwide Pharmaceutical Markets -- 1960 to 1999," *Drug Information Journal*, Jan-Mar 2001.

³ As $Npq > 10$, we use the normal distribution to approximate the binomial.

lead to death or serious injury, the FDA may require that the label warning be presented in a black box.⁴ These Black Box Warnings, or BBWs, can be imposed either before or after a drug has been approved for marketing, and they present important safety information uncovered during the drug review or after the drug has been on the market for some period of time. By defining a “safety event” as the addition of a BBW to the label, instead of as a safety withdrawal, one gains considerable sample size and statistical power.⁵ It is important to emphasize, however, that using a BBW to define a “safety event” still captures severe safety problems because a BBW identifies drugs with adverse effects that may potentially lead to death or serious injury. Moreover, because BBWs can be applied prior to drug marketing, they may be used to evaluate the safety assessment carried out by the FDA during drug review, something we explicitly consider in the current paper. Indeed, the recent IOM report recommends that the FDA report the timeliness of its regulatory actions, including black boxes added to labels.⁶ Timely identification, confirmation, and communication of risks and benefits are the best measure of regulatory success.

Our paper will proceed as follows. Section II will describe the unique dataset we employ in our analyses to study the determinants and timing of BBWs. This section will also describe our model covariates and the rationale for their inclusion in the data sample. We believe these data represent a significant improvement over the frequently used withdrawal data, and may provide an excellent opportunity to study trends in drug safety

⁴ The regulatory basis of the black box warning is at 21 C.F.R. § 201.57(e).

⁵ For our analysis, we treat a drug withdrawn for safety reasons as if it also receives a BBW at that time.

⁶ Recommendations to report BBW information and other ways of improving post marketing risk communication activities are on Page 3-27 of the of Institute of Medicine, *The Future of Drug Safety: Promoting and Protecting the Health of the Public*. Washington, DC: National Academies Press, 2006.

over time and across therapeutic drug classes. The potential applications of BBWs as a proxy for drug safety are wide ranging.

In Section III we present and empirically test several hypotheses. The two principal hypotheses of the section, however, concern the impact of the Prescription Drug User Fee Act (PDUFA) of 1992, and its subsequent renewals in 1997 and 2002, on drug safety, as measured by BBWs. Under PDUFA, manufacturers pay the FDA user fees to hire additional staff to review new drug applications (NDAs) and expedite approval times, which prior to the Act were growing steadily over time, and averaged over two years in 1992.⁷ While recent economic research has documented the significant social benefits of PDUFA via improved access to new drugs and increased incentives to invest in research & development (R&D),⁸ critics of PDUFA have raised concerns over drug safety because of the shorter FDA review times.⁹ For this reason, one of the principal aims of our analyses will be to examine the impact PDUFA may have had on drug safety, as measured by the prevalence and timing of BBWs. Section IV will conclude the paper and make suggestions for future research.

II. DATA AND MODEL VARIABLES

This study employs a highly unique FDA dataset, one that is based on twenty-five years of historical BBWs data. Specifically, we have tracked therapeutic NMEs

⁷ Joseph A. DiMasi (2001) "New Drug Development in the United States from 1963 to 1999," Clinical Pharmacology and Therapeutics 69, 286-296.

⁸ Tomas J. Philipson, Ernst R. Berndt, Adrian H.B. Gottschalk, and Matthew W. Strobeck (2005), "Assessing the Safety and Efficacy of the FDA: The Case of the Prescription Drug User Fee Acts," NBER Working Paper 11724, Cambridge, Massachusetts; John A. Vernon, Joseph H. Golec, Randall W. Lutter, and Clark Nardinelli (2006), "FDA New Drug Approval Times, Prescription Drug User Fees, and R&D Spending," AEI-Brookings Joint Center for Regulatory Studies. Available at: <http://www.aei.brookings.org/publications/index.php?tab=topics&topicid=44>.

submitted to the FDA between May 4, 1981 and May 31, 2005 that were subsequently approved and marketed.¹⁰ For these 516 NMEs, we included BBWs required in the labeling through February 16, 2006.¹¹ We used MedWatch to identify instances of safety-related labeling changes.¹²

Our database also includes several drug-specific attributes that we hypothesize, *a priori*, will influence the likelihood of receiving a BBW. First, we record the review status of drug applications received by the FDA. Drugs that are novel therapies are granted priority status by the FDA, and as a result their NDA receives a faster review. Because the therapy is novel, and because an expedited review may be more likely to miss some potential safety problems, priority drugs might have higher rates of post-market BBWs. Some previous empirical research has suggested a link between review times and post-market adverse drug reactions.¹³ Second, some NMEs are for rare, or “orphan,” diseases. These drugs generally have limited clinical testing prior to approval, and therefore they may have a higher rate of post-market BBWs. Third, drugs primarily used in hospital settings, which are frequently taken by the seriously ill, might have lower rates of BBWs because these drugs typically have different safety profiles than drugs

⁹ Public Citizen Health Research Group (2002), “Comments on the Prescription Drug User Fee Act (PDUFA)” Docket No. 01N-0450, January 25, 2002. Available at: <http://www.citizen.org/publications/>.

¹⁰ The starting date was chosen initially to create a data set to equalize the length of time before and after the enactment of PDUFA. The data set has since been updated, so the significance of the start date no longer exists. The term “therapeutic” means that we are excluding NMEs that do not treat disease, such as diagnostic products.

¹¹ In some instances, a single BBW decision will apply to an entire class of drugs. Here, we count each NME receiving a BBW as a separate incident, but we have collected some information on “class labeling” and include it as Appendix A.

¹² MedWatch is FDA’s Safety Information and Adverse Event Reporting Program. For more information, see <http://www.fda.gov/medwatch/What.htm>, accessed 7/7/2006.

¹³ Mary K. Olsen (2002), “Pharmaceutical Policy Change and the Safety of New Drugs,” *Journal of Law and Economics* 45:615-642.

purchased through pharmacies and taken at home.¹⁴ Fourth, we use a measure of the first year prescription volume derived from IMSTM and VerispanTM data because some safety issues are not revealed until there has been sufficient experience with a particular drug. All else held constant, safety issues will be identified faster among drugs with higher initial prescription volume. Finally, the drug review process itself changed with the enactment of PDUFA in 1992. User fees funded the acceleration of the review process for NDAs. Some of the criticisms of PDUFA suggest that the faster initial review could, in theory, increase the rate of post-market BBWs. Moreover, under both PDUFA I and PDUFA II (but not PDUFA III), user fees could not be used for post-market safety surveillance. For the statistical analysis, we define *PDUFA* as a binary variable that takes on the value of unity if the drug was approved post-PDUFA, and zero otherwise.

We distinguish between BBWs added to drug labeling during the drug review process, “pre-market BBWs,” and those applied after a drug has already been approved and marketed, “post-market BBWs.” Post-market BBWs are of particular importance because they represent instances where there is an acknowledgement that a drug was initially marketed without all of the information later deemed necessary for safe and effective use. We summarize these data in Table 1.

¹⁴ Although we refer to such NMEs “Hospital Drugs,” the moniker identifies primary use in all non-retail settings, including hospitals, nursing homes, and clinics.

Table 1: Characteristics of Marketed Therapeutic NMEs May 1981- May 2005

	NMEs	Priority Review	Hospital Setting	Orphan Status	Pre-market BBW
Pre-PDUFA	228	116	44	40	30
PDUFA	288	116	21	46	51
Total	516	232	65	86	81

III. RESEARCH METHODS AND PRINCIPAL HYPOTHESES

In the current analyses, we will employ three methods to study several different hypotheses involving changes in the rate of BBWs over time. First, we examine the rates of BBWs applied before a drug is marketed both before and after the enactment of PDUFA. By directly comparing these rates, we can analyze the effects during the review of the drug.

Second, we will analyze the rates of post-market BBWs. We use the Kaplan-Meier methodology and test for a difference in the BBWs rate for drugs submitted to the FDA before and after PDUFA. We supplement this simple, exploratory approach with a more thorough multivariate analysis, which includes covariates that control for other factors hypothesized to affect the rate of post-market BBWs. These multivariate analyses may shed light on key determinants of BBWs.

Finally, we undertake an analysis of post-market BBWs accounting for analytic problems associated with using submission cohorts to analyze surveillance. That is, to accurately assess the impact of safety surveillance practices instituted by PDUFA, we

develop a model that controls for the length of time a drug has been marketed under the PDUFA regime, regardless of whether it was approved after the enactment of PDUFA.

While we test several hypotheses on the determinants of BBWs in this paper, we explicitly define the two that concern the possible influence of PDUFA on the incidence and timing of BBWs. From a policy perspective, these hypotheses are the most germane to the ongoing debate over the social costs and benefits associated with PDUFA. We turn now to our first principal PDUFA-related hypotheses.

Hypothesis 1: *PDUFA Accompanied a Change in the Rate of Pre-market BBWs*

Black Box Warnings are often applied during the review of the application, prior to marketing. In our sample, there were 81 pre-market BBWs, 30 of them occurred pre-PDUFA and 51 since its enactment in 1992. One might expect the expedited reviews under PDUFA to lead to fewer pre-market BBWs, but as shown in Table 1, the fraction of NMEs given pre-market BBWs after the enactment of PDUFA is actually higher, 0.177 versus 0.132. Of course, it is not immediately apparent how much of this difference is attributable to PDUFA and how much is attributable to other factors that changed over the sample time period. Table 2 summarizes the distribution of the 516 NMEs in our sample by review status and PDUFA period.

Table 2: Comparison of the Incidence of Premarket BBWs in pre-PDUFA and PDUFA NME Labeling by Review Type

Among NMEs Lacking Premarket BBWs			
Review Status	Pre-PDUFA	PDUFA	Totals
TOTAL	198	237	435
Standard	107	151	258
Priority	91	86	177

Among NMEs with Premarket BBWs			
Review Status	Pre-PDUFA	PDUFA	Totals
TOTAL	30	51	81
Standard	5	21	26
Priority	25	30	55

In assessing the associations among pre-market BBWs, review status (standard or priority), and PDUFA era (before and during), we use a simple cross-tab analysis of a $2 \times 2 \times 2$ contingency table. This approach enables us to determine which combination of variables best explains the variation observed in Table 2, while minimizing the number of interactions required to account for the variation. We find that pre-market BBWs are associated with review status; review status is associated with PDUFA; and pre-market BBWs are associated with PDUFA. These associations, however, were not statistically significant at the 0.05-level. There was a higher rate of pre-market BBWs for priority reviews than for standard reviews, and there were relatively more priority reviews pre-PDUFA; both of these associations were statistically significant at the 0.05-level. Pre-market BBWs appear to be more common under PDUFA, but this difference fails a conventional test for statistical significance ($p = .051$).

When controlling for a drug's PDUFA status, pre-market BBWs are 34 percent more likely among priority review NMEs than they are for those receiving standard reviews. Controlling for whether an NME receives a priority or standard review, pre-

market BBWs are 13 percent more likely to occur if a NME was approved after PDUFA than if it was approved before. Standard reviews are 13 percent more likely under PDUFA than they were pre-PDUFA. Lastly, under PDUFA the rate of pre-market BBWs for standard review NMEs is much higher than they were prior to PDUFA (12.2 percent vs. 4.5 percent).

The statistical association between PDUFA and the rate of pre-market BBWs is, however, potentially misleading. The likelihood of a drug receiving a pre-market BBW is affected by a drug's review status, and its review status is affected by whether or not it was approved before the enactment of PDUFA. After accounting for review status, the association between PDUFA and the rate of BBWs remains positive, but is no longer statistically significant at the 0.05-level. This exploratory analysis, therefore, does not find empirical support for the hypothesis that PDUFA accompanied a statistically significant change in the rate of pre-market BBWs. Future researchers, however, may wish to probe this association more closely by explicitly considering the changing impact of PDUFA on drug approval times over time. User fees have grown considerably since 1992, and reductions in FDA approval times were not instantaneous.¹⁵ We next study whether or not PDUFA influenced the rate of post-market BBWs.

Hypothesis 2: *PDUFA Accompanied a Change in the Rate or Timing of Post-market BBWs*

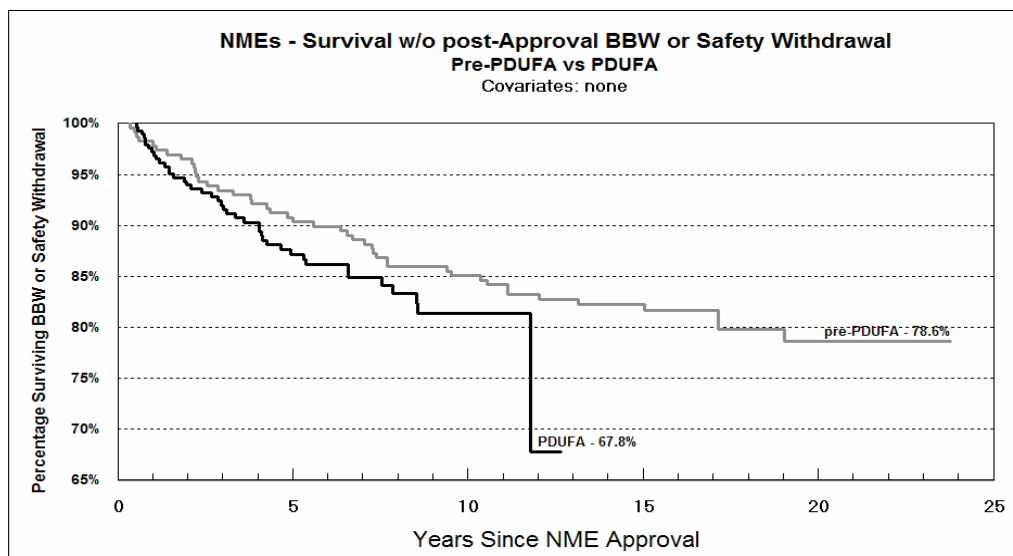
As previously mentioned, the comparison of rates of post-market BBWs is complicated by the fact that different drugs have been on the market for different lengths

¹⁵ For an excellent analysis of the gradual impact of PDUFA on FDA approval times since 1992, see Ernst R. Berndt, Adrian H. B. Gottschalk, Tomas J. Philipson, and Matthew W. Strobeck (2004) "Assessing the Impacts of

of time. These different life spans, or vintages of drug cohorts, make comparisons between groups of drugs from different eras difficult. Because of this difference, we analyze post-market BBWs using a survival analysis methodology. Survival analysis, using a Kaplan-Meier estimator, allows us to capture events occurring through time, and then directly compare these relative rates across drug cohorts.

We ignore pre-market BBWs, and include only the first BBW applied after a drug has been on the market. Once the BBW event occurs, the NME is removed from the analysis. We divide the set of NMEs according to whether they were approved before or after the enactment of PDUFA. The Kaplan-Meier graphs for these two groups are shown below in Figure 1.

Figure 1: Survival of NMEs by PDUFA Status



Fully two-thirds (68 percent) of PDUFA-era therapeutic NMEs survived nearly 13 years without any new BBWs (or safety withdrawals). Among the pre-PDUFA NMEs, 79 percent survived to the end of the data sample period without new BBWs. The shapes and

locations of the PDUFA and pre-PDUFA curves are similar until year 12. The curves diverge at year 12, but there are so few underlying observations that this divergence is of only marginal importance. There are only six NMEs remaining in the PDUFA cohort at the time of the safety event occurring in year 12.

To test the equality of the two survival curves, we used a Tarone-Ware weighted log rank test, which gives greater weight to the sections of the curves with more underlying observations. We failed to reject the null hypothesis of equality ($p = 0.22$). The curves are not statistically different from one another.

Because the faster reviews and greater resources provided by PDUFA represent only one of several factors that may influence post-marketing BBWs, we added to our analysis four drug-specific attributes that plausibly might affect the rate of post-market BBWs. These covariates are i) first year prescription volume; ii) whether a drug is typically distributed through hospitals; iii) whether the drug received a priority review; and iv) whether the drug was approved through the orphan drug program.

Theoretically, we would expect drugs with greater initial patient exposure to be associated with a faster identification of safety issues, and thus a higher likelihood of receiving a BBW. It also seems reasonable to expect that faster FDA reviews, which are given to novel therapies receiving priority review status, might lead to a higher hazard of a post-market BBW. Safety problems among drugs distributed through hospitals (as opposed to those being distributed through pharmacies) might be difficult to identify because the individuals taking these drug tend to be already infirm. Finally, orphan drugs, which are developed for rare diseases, and which typically have limited patient populations, might be associated with a different post-market hazard rate for BBWs.

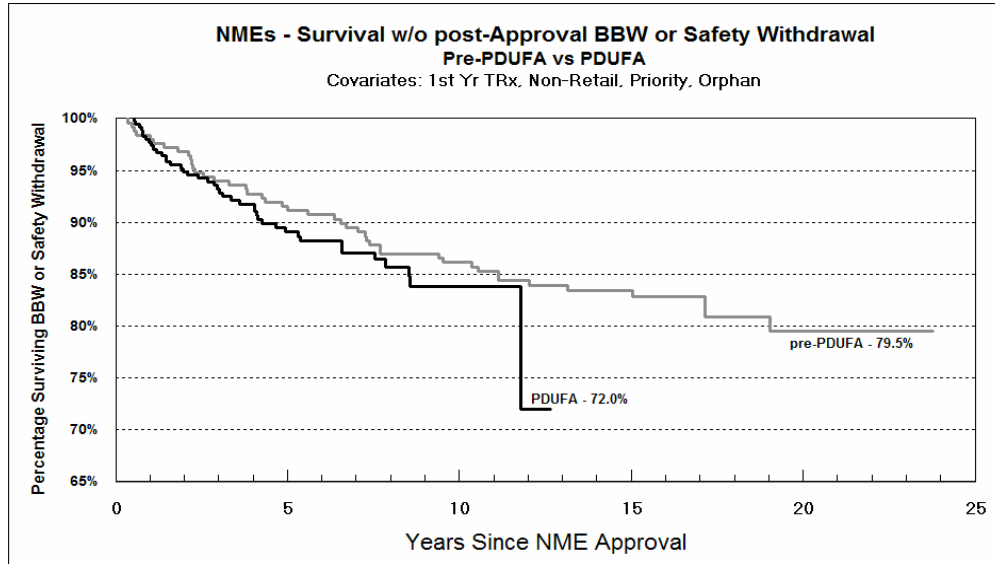
To more rigorously test our second hypothesis, we use a stratified version of the Cox Proportional Hazard Model and include as covariates the four aforementioned NME-specific attributes hypothesized to influence the hazard rate for BBWs. In stratifying by PDUFA status, we allow each group of NMEs to have its own baseline hazard, but constrain the relative hazard to be the same across both strata. More specifically, we use the semi-parametric Cox Proportional Hazard Model to measure the extent to which a covariate shifts a hazard function up or down. The Cox model is flexible in that it does not require an explicit specification of the model's functional form. In a Cox Proportional Hazard Model regression model, a hazard function for observation j is defined by the following:

$$h_j(t) = h_0(t)e^{x_j\beta} = h_0(t)\exp(b_1x_{1j} + b_2x_{2j} + \dots + b_kx_{kj}) \quad (1)$$

The first of the two components appearing in equation (1), $h_0(t)$, is the *baseline hazard*. It is not parameterized and is not estimated. The other component, $e^{x_j\beta}$, is sometimes called the *relative hazard*, and it captures the effect of the observed characteristics on the hazard function. The effect of the vector of covariates, \mathbf{X} , on the hazard rate is independent of time, t . The hazard for one observation is proportional to the hazard for another observation and this proportionality constant is independent of time (hence the name “proportional hazard model”). By exponentiation with the model coefficient estimates, we obtain hazard ratios for a one-unit change in the values of the model covariates. For example, a coefficient estimate of 2.0 implies that a one-unit change in the covariate would double the hazard rate. Estimating equation (1) using the four aforementioned covariates (drug characteristics) we find that ultimate survival

without a post-market BBW is 72 percent for the PDUFA cohort and 80 percent for the cohort approved before PDUFA. This is depicted below in Figure 2.

Figure 2: Survival of NMEs by PDUFA Status with Drug-Specific Covariates



Our Cox regression results from this model are summarized below in Table 3. We present the hazard ratios, standard errors, t-statistics, and p-values for each model covariate.

Table 3: Hazard Ratio Analysis of Safety Events

Variable	Hazard Ratio	Standard Error	t-Statistic	p-value
First Year Rx Volume	1.1316	0.0490	2.8538	0.0043
Hospital Drug	0.2691	0.1399	-2.5240	0.0116
Priority Review	1.6431	0.3718	2.1941	0.0282
Orphan Drug	0.5979	0.2213	-1.3896	0.1648

As hypothesized, initial prescription volume is associated with an increase in the hazard rate of a BBW, and this result is significant at 0.01-level. Also as expected, hospital drugs are associated with lower hazard rates. This result is significant at 0.05-level. This lower hazard rate for hospital drugs may be the result of the fact that the BBW system is focused on drugs distributed through retail channels and taken at home; the difficulty of identifying adverse drug events in a population already sick may also contribute to the low hazard rate. In accordance with expectations, we also find that drugs given a priority review have a higher hazard rate, and this is statistically significant at the 0.05-level. Lastly, while the coefficient on the orphan drug dummy variable carries the theoretically anticipated negative sign, it is not statistically significant at conventional levels.

Including the drug-specific attributes in our model, not surprisingly, results in the two survival curves appearing closer together; in fact, they almost touch in year 8 and in year 11. The close proximity of the curves in Figure 2, along with the statistically significant results reported in Table 3, suggests that much of any difference in timing and incidence of BBWs before and after the enactment of PDUFA is associated with the attributes of the drugs reviewed under each regime. While the evidence thus far appears to suggest that PDUFA was not accompanied by a change in the rate of post-market BBWs, it is possible that unmeasured intertemporal variability in the effectiveness of the PDUFA surveillance regime may be partially influencing our results.

In particular, review status may not fully capture differences in surveillance across regimes. Drugs approved after the implementation of PDUFA spend their entire marketing lives under the PDUFA surveillance program. Drugs approved shortly before

the implementation of PDUFA would spend almost all of their observed lives under PDUFA. Drugs approved near the beginning of our pre-PDUFA cohort would spend about half their market lives under PDUFA. Clearly then, if the surveillance program became more effective after the passage of PDUFA, we would see lower hazard rates among drugs approved in the years just after PDUFA was enacted.

To control for this effect, we add to our analysis the variable *Months Pre PDUFA*, which is equal to the number of days divided by 30 between approval and September 1, 1992. For drugs approved after September 1, 1992, the variable is set to zero. We were concerned that this variable, created to capture the effects in drugs approved prior to PDUFA, would simply pick up some overall time trend. Therefore, we have also added to our model the variable *Months Pre & Post PDUFA*. This variable is measured as the number of days divided by 30 between approval and September 1, 1992. It is the same as *Months Pre PDUFA* for drugs approved prior to PDUFA. A drug approved after PDUFA would have a negative value for *Months Pre & Post PDUFA*, but would have a zero for *Months Pre PDUFA*. Thus, the coefficient for *Months Pre PDUFA* measures the *proportional* increase in the hazard function given one additional month a drug is marketed prior to PDUFA.

We define the dummy variable *PDUFA*, which is set equal to one for drugs submitted for approval on or after September 1, 1992 and zero otherwise. The results using *PDUFA* alone in the Cox model are reported in column 1 of Table 4. This coefficient is not statistically significant at conventional confidence levels, consistent with our results reported previously in Table 3.¹⁶

¹⁶ The coefficient from our analysis tells us whether the hazard ratio is significantly different from unity. We report in Tables 3 and 4 the hazard ratio and the standard error of the coefficient. A value of 1.3355

Table 4: Hazard Ratio Analysis of Safety Events

(Coefficients are estimated hazard ratios; standard errors are in parentheses)

Variable	(1)	(2)	(3)	(4)	(5)
PDUFA	1.3355 (0.3109)	1.3924 (0.3255)	1.2716 (0.2984)	0.8639 (0.3393)	0.8416 (0.3361)
Priority	--	1.4551 * (0.3173)	1.7300 ** (0.3894)	1.8457 ** (0.4174)	1.7730 ** (0.4041)
Orphan	--	--	0.5300 * (0.1936)	0.4697 ** (0.1731)	0.5347 * (0.2001)
Hospital	--	--	0.2681 ** (0.1394)	0.3767 * (0.2034)	0.3870 * (0.2100)
Months Pre PDUFA	--	--	--	0.9889 (0.0073)	0.9888 (0.0074)
Months Pre & Post PDUFA	--	--	--	0.9998 (0.0049)	1.0000 (0.0050)
First Year Prescriptions	--	--	--	--	1.1297 ** (0.0487)

* Significant at the 10% confidence level

** Significant at the 5% confidence level

Since before the start of our sample period, the FDA has been giving drugs priority status designations based on whether they presented a novel therapeutic benefit.¹⁷ A change in the balance of priority-reviewed drugs may have an effect on the overall likelihood of a BBW. This is because these drugs, with their ability to provide a novel therapeutic benefit, may also involve unusual risk. We account for this change in the balance by including the model variable *Priority*, which is set equal to one for priority drugs and zero otherwise. These results appear in column 2 of Table 4. Priority drugs have a higher hazard of a safety event, which is not unexpected. Again, the regime variable *PDUFA* is not significant at the tested confidence levels.

implies that the variable is associated with 34% increase in the hazard rate. The coefficient is the log of the ratio, or $\ln(1.3355) = 0.2893$. As the standard error is 0.3109, the coefficient on *PDUFA* is not significant at the 90% confidence level.

We also wanted to correct for several other potential confounding effects. “Orphan” drugs are given special review status because they are intended for small patient populations. The orphan drug review is generally restricted to small clinical trials. Moreover, the small number of patients treated with orphan drugs makes it difficult to monitor drug safety after marketing. Therefore, the binary variable *Orphan* is defined to be equal to one for orphan drugs and zero otherwise. As shown in column 3 of Table 4, *Orphan* drugs have lower hazards of safety events. The small patient populations may make surveillance more difficult and safety events less likely. The variable *Priority* is, again, associated with a higher hazard rate, and *PDUFA* continues to be statistically insignificant.

The effects of adding *Months Pre PDUFA* and *Months Pre & Post PDUFA* are shown in column 4. Neither of these variables is found to be statistically significant. Column 5 includes the addition of the first year prescription variable, which is significant at the 0.05-level. As was reported in Table 3, a million additional prescriptions in the first year on the market is associated with a 13% increase in the rate of BBWs. *Months Pre PDUFA* is not statistically significant, but *Orphan* and *Hospital* are, and are associated with, as hypothesized, a lower rate of BBWs. The variable for priority drugs suggests a higher rate of BBWs; this finding is significant at 0.05-level.

In sum, our regression analyses generate several results worth emphasizing. Drugs with higher initial prescription volume have a higher rate of postmarket BBWs, supporting the hypothesis that relatively rare side effects are discovered earlier when there is a larger initial patient population. Also, drugs receiving priority review have

¹⁷ Since the implementation of PDUFA, FDA switched from a review system where drugs were prioritized as being A, B or C to simply P (priority) or S (standard). We assume pre-PDUFA NMEs designated A or B

higher rates of BBWs, which could be caused by the relatively smaller initial clinical studies, the fact that the therapy is more novel and thus less well understood, or both of these factors. Drugs sold through hospitals have lower hazard rates, possibly because it is difficult to identify safety risks when the patient population is already very sick. Orphan status, which we found to be statistically significant at the 0.10-level, is also associated with a lower hazard level, perhaps because of the smaller patient populations.

Regarding our second principal hypothesis, we find no statistically significant difference in the rate of postmarket BBWs after PDUFA. The comparable rates of survival just prior to the last PDUFA failure (occurring at 11.8 years) are 84.4 percent for pre-PDUFA NMEs and 83.9 percent for PDUFA NMEs. This is important because critics of PDUFA have raised concerns of compromised drug safety, and while our findings in this section are by no means definitive, they do suggest that safety has not been adversely affected as a result of PDUFA.

IV. CONCLUSION

This research was motivated by the inadequacy of using the percentage of safety withdrawals as a way to quantify drug safety in the U.S. Safety withdrawals capture the most extreme events, they are influenced by non-safety factors, and they cannot be subject to rigorous statistical analyses. The use of percentages can be misleading because it does not account for the length of time a drug has been on the market. Analyzing a less severe, but more common safety event, such as the application of a Black Box Warning, addresses these shortcomings with safety withdrawals. Not only are drugs with BBWs a

to be equivalent to those currently designated P.

better reflection of the universe of drugs, they can also be used in a variety of types of statistical analyses, as we have demonstrated in this paper. We have also demonstrated the use of survival analysis as a method to compare rates of BBWs when drugs have been on the market for different lengths of time.

As expected, we find that the rate of application of BBWs depends on factors such as whether a drug received a priority review, its initial prescribing volume, and whether it was distributed primarily in hospital-type settings. We are unable to identify statistically significant trends in the rates BBWs over time. This is an important finding because PDUFA has been criticized on the grounds that it may compromise drug safety. We find no evidence of this in the current study. Finally, given that recent economic research suggests PDUFA has imparted significant benefits to society through improved access to new drugs and greater incentives for R&D, our research suggests, on net, PDUFA is likely to be generating social benefits in excess of costs. Of course, further research is warranted, and we have made several suggestions for future directions.

REFERENCES

Berndt, E, Gottschalk, A. Philipson, T and Strobeck, M (2004) “Assessing the Impacts of the Prescription Drug User Fee Acts (PDUFA) on the FDA Approval Process,” *AEI-Brookings Joint Center for Regulatory Studies*.

DiMasi, JA (2001) “New Drug Development in the United States from 1963 to 1999,” *Clinical Pharmacology and Therapeutics* 69, 286-296.

Fung, M. (2001) “Evaluation of the Characteristics of Safety Withdrawal of Prescription Drugs from Worldwide Pharmaceutical Markets -- 1960 to 1999,” *Drug Information Journal*, Jan-Mar 2001.

Institute of Medicine (2006) *The Future of Drug Safety: Promoting and Protecting the Health of the Public*. Washington, DC: National Academies Press, 2006.

Olsen, MK (2002), “Pharmaceutical Policy Change and the Safety of New Drugs,” *Journal of Law and Economics* 45:615-642.

Philipson, T. Berndt, E. Gottschalk, A. and Strobeck, M. (2005), “Assessing the Safety and Efficacy of the FDA: The Case of the Prescription Drug User Fee Acts,” *NBER Working Paper Series*, Working paper 11724, Cambridge, Massachusetts.

Public Citizen Health Research Group (2002), “Comments on the Prescription Drug User Fee Act (PDUFA)” Docket No. 01N-0450, January 25, 2002. Available at: <http://www.citizen.org/publications/>.

Vernon, J. Golec, J. Lutter, R. Nardinelli, C. (2006), “FDA New Drug Approval Times, Prescription Drug User Fees, and R&D Spending,” *AEI-Brookings Joint Center for Regulatory Studies*.