PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Postmarket Safety in Canada: Are Significant Therapeutic Advances and Biologics Less Safe Than Other Drugs? A Cohort Study
AUTHORS	Lexchin, Joel

VERSION 1 - REVIEW

REVIEWER	Thomas J Moore Institute for Safe Medication Practices, United States
REVIEW RETURNED	11-Nov-2013

GENERAL COMMENTS	The author is a well-respected Canadian researcher with numerous previous publications on pharmaceutical policy, drug efficacy and safety. The current submission builds on and expands on data from a 2012 study published in the Archives of Internal Medicine (now JAMA Internal Medicine). While the current study contains valuable systematic assessment of Canadian post approval regulatory decisions over a substantial period of time, unfortunately I regard it as unsuitable for publication in this form. My reasoning is as follows: A statistical study based on “proving” rather than disproving the null hypothesis succumbs to Type II statistical error unless accompanied by a substantial analysis capable of identifying the characteristics of the differences, should they exist. Otherwise it is impossible to tell whether the study truly concludes there is no difference, or that the methodology was sufficiently flawed that it could not detect a true difference. In drug bioequivalence assays and non-inferiority clinical trials, for example, a finding of equivalence or non-inferiority is supported by such analysis, although even these methods are not without controversy. These studies also contain validated endpoints. The next issue concerns the conclusions reached that there is “no difference in safety” of biological products compared to small molecule drugs. (p13/line 25). Next at line 32 the finding is described as “postmarket safety,” a different concept. In the next paragraph it is further expanded to say “the safety profile of NAS is the same regardless of the level of therapeutic benefit.” Part of this is a language problem. The investigator has observed post-market safety warning actions by a regulator. That is not a proxy for “safety” or a “safety profile,” and I suggest they should not be considered synonyms. The safety profile of a drug needs to include any boxed warnings, contraindications and other risks detected prior to approval as well as post approval. The second part of the problem is the limitations of the null hypothesis

	The third limitation with the safety warning metric is the wide variety in the incidence and severity of safety problems involved. A warning might indicate a handful of reported cases of a catastrophic side effect (efalizumab, progressive multifocal leukoencephalopathy), a problem for a small patient subset (fluoroquinolones, worsening of myasthenia gravis), or a massive contraindication (dolasetron – any therapeutic use under 18 years age, any post operative nausea). The comparison involving therapeutic advances also appear to be definition sensitive. In the Archives study the author used a priority review as the primary index of therapeutic value and reported a substantial and statistically significant difference between priority and standard review drugs and subsequent safety warnings. In the current study, a different definition of therapeutic value was created using two versions of a Canadian formulary review board criteria for some agents and recode of ratings from the French journal Prescrire for others. In the new definition the differences seen in the previous study do not reach statistical significance. I'm somewhat puzzled by the therapeutic value issue. Given a subjective judgment, it does not seem surprising to discover differences, particularly with different criteria for designating a novel drug. And using 3 different rating schemes for one variable does not help, and is probably not the best solution to addressing the missing value problem. Also, some distortion is introduced by using the median time to first safety warning, rather than the median and range for all the serious safety warnings. The latter seems to be a more accurate index of when prescribing physicians will learn about drug risks not detected or fully understood at time of approval. This complicates a survival analysis, but measuring all warnings probably provides a more accurate measure of how long after approval it takes to detect substantial new safety issues. In two published studies in the US the median time to all warnings was 11 years (compared to 4 years to first warning in these data). Arch Intern Med. 2012 Jan 9;172(1):78-80. doi: 10.1001/archinternmed.2011.618. The FDA and new safety warnings. Moore TJ, Singh S, Furberg CD. Pharmacoepidemiol Drug Saf. 2013 Mar;22(3):302-5. doi: 10.1002/pds.3395. Epub 2013 Jan 2. Evaluation of FDA safety-related drug label changes in 2010. Lester J, Neyarapally GA, Lipowski E, Graham CF, Hall M, Dal Pan G.

REVIEWER	Nigel S B Rawson Eastlake Research Group, Oakville, ON, Canada and Fraser Institute, Vancouver, BC, Canada I have published work in the same area previously. I was employed by GlaxoSmithKline Canada between June 2004 and March 2012.
REVIEW RETURNED	14-Nov-2013

GENERAL COMMENTS	I believe that the methods are not comprehensive enough, appropriate references are omitted, and the limitations of the methods need to include comparison with other work ***Introduction*** Dr Lexchin begins his article by stating that drug safety is becoming a topic of increasing concern in Canada. However, since the thalidomide tragedy, drug safety has been of ongoing concern to Canadians, which increases periodically when drugs, such as practolol, benoxaprofen, troglitazone, cisapride, cerivastatin and rofecoxib, taken by large numbers of patients are withdrawn for safety reasons. What is of greater concern is how Health Canada identifies drug safety issues, with only an under-funded voluntary adverse drug reaction (ADR) reporting system[1] that can in no way be regarded as a reliable safety net, and the frequent lack of clarity and timeliness in its communications about drug risks.[2] Dr Lexchin also states that ADRs "are the fourth to sixth leading annual cause of death." This assertion is based on a 1998 meta-analysis[3] that was criticized at the time of its publication as significantly over-estimating the number of fatal ADRs due to using studies too diverse to aggregate.[4,5] Moreover, 60% of the studies were published before 1980 and had higher incidences of ADRs, whereas the later studies had much lower incidences.[4] A repeat meta-analysis using the same methods performed in 2000 by experts in the technique demonstrated that ADRs were not defined with sufficient precision and consistency across the studies to permit the pooling of results, correct numbers were not used for calculating their incidence rates, and appropriate adjustments were not made for differences in studies, patients and outcomes, which likely led to over-estimation of the number of patients affected.[6] It is time for this urban myth to be abandoned.[7] ***Methods*** Dr Lexchin uses data on 406 new drugs approved in Canada between January 1997 and March 2012 obtained from Health Canada's annual performance reports and links them with information on serious safety warnings from the agency's MedEffect website "supplemented by searching on the product name in the Drug Product Database (DPD)". The DPD allows access to the relevant Product Monographs (PMs), which are the definitive legal documents on drugs in Canada and include all warnings (serious or not). Serious warnings are usually highlighted by a box but, in older PMs, bold capital letters alone may be used. Dr Lexchin groups the drugs using the Patented Medicine Prices Review Board classification of new drugs, which is an exceedingly restrictive categorization such that few products are considered to be breakthroughs or substantial improvements. If a drug for the disease

	or condition exists, whether it is highly or poorly effective, it is rare for a new, more effective drug to be considered as a breakthrough or substantial improvement. Dr Lexchin assesses, as his principal objectives, the survival rates of the drugs until they acquire a serious safety warning or are discontinued for a safety reason, comparing biologic versus traditional drugs (although the performance reports "only gave the type of product … from 2000 onwards") and drugs with "significant therapeutic advances" versus those without. ***Results*** Dr Lexchin finds that, of 406 new drugs approved between January 1997 and March 2012, 87 (21.4%) acquired a serious safety warning or were discontinued for a safety reason. This is similar to the rate of 19.4% that he reported elsewhere when he examined, using the same methods, 434 new drugs approved in Canada between January 1995 and December 2010.[8] However, these figures are differ substantially from those I recently reported based on a study of the approval times and the acquisition of a serious safety warning or discontinuation for a safety reason for 454 new therapeutic (i.e. new salts, esters, dosage forms and combinations of previously approved drugs, diagnostic products and vaccines were excluded – it is unclear whether Lexchin had any exclusions) approved between 1992 and 2011 in both Canada and the United States.[9] Drugs discontinued for safety reasons were identified from websites of Health Canada and US Food and Drug Administration (FDA) and publications. Serious safety warnings in Canada were identified from the relevant PM obtained from the DPD website, supplemented with information from the MedEffect website. Black-box warnings (BBWs) were obtained from the history of each drug's Label available on the FDA's website, supplemented with information from the FDA's MedWatch website. In contrast to Dr Lexchin's results, I found that 199 (43.8%) of the new drugs acquired a serious safety warning or were discontinued for a safety reason in Canada, while 158 (34.8%) acquired a BBW or were discontinued for a safety reason in the United States. This difference in the rate of acquisition of serious safety warnings leads to substantial differences in the probability of acquiring a serious safety warning or safety discontinuation (Table 1) between biologics and traditional drugs. Lexchin found no statistical difference between the two types of drugs (log-rank test, $p=0.47$), whereas I found significant differences between them in both Canada and the United States (both log-rank tests, $p<0.0001$). Table 1: Differences in the probability of acquiring a serious safety

warning or being discontinued for safety

	Lexchin		Rawson[9]			
Data source	MedEffect		Product Monograph plus MedEffect		Label plus MedWatch	
Time period	Jan 1997 – Mar 2012		Jan 1992 – Dec 2011		Jan 1992 – Dec 2011	
Survival horizon	Unspecified		10 years		10 years	
	No.	Survival rate	No.	Survival rate	No.	Survival rate
Biologics	79	27.3%	60	70.9%	60	55.1%
Traditional drugs	219	29.9%	394	37.2%	394	27.3%

Why the difference in results? I believe that the use of PMs, supplemented by MedEffect information, is a more comprehensive assessment of the acquisition of serious safety warnings than the use of MedEffect alone. MedEffect frequently omits serious drug warnings. The Canadian data using PMs (plus MedEffect) are also consistent with US data using Labels (plus MedWatch), although the Canadian rates are higher. Dr Lexchin's article is deficient without some discussion of the difference between his methods and results and mine and the limitations imposed by relying mainly on MedEffect.

The title of Dr Lexchin's article seems to imply that an increase in the use of serious safety warnings would indicate that unsafe drugs are being approved. However, safety warnings are exactly that – *warnings*. They are not indicative of actual health outcomes. The only solid evidence about overall drug safety is the rate of discontinuation for safety reasons, which in drugs approved in 2002-2011 was 2.1% in Canada and 1.3% in the United States, which was marginally less than the corresponding rates of 2.3% and 1.7%, respectively, in 1992-2001.[9]

Drugs are not approved by regulators unless their benefits outweigh their risks. Health Canada and the FDA have been working to improve the timeliness of the approval of innovative drugs in order to get them as quickly as is safely possible to patients. One way that they have done this is by increasing the number of drugs receiving a

priority status review. Many of these products are innovative pharmacologically-complex drugs that may provide novel benefits but may also have unusual risks. To ensure that patients and healthcare providers are aware of this, Health Canada and the FDA have increasingly required serious warnings in the PMs and Labels over the last decade, frequently at the time of approval.[9] An increase in their use does not necessarily mean that unsafe drugs have been approved.[10]

Other potential reasons exist for the increased use of serious warnings. For example, following several high-profile discontinuations due to safety reasons between 1999 and 2004, regulatory agencies may simply be applying greater caution. Another reason could be that regulatory agencies are trying to provide more relevant information to allow patients and healthcare providers to make more informed decisions about the use of new drugs. In addition, the greater issuance of serious safety warnings in Canada relative to the United States could be a response to criticism received by Health Canada regarding the timeliness and adequacy of its risk communications following a well-publicized coroner's inquest into the death of a teenager while taking cisapride.[11]

Only time will tell whether Health Canada is more prescient than the FDA in predicting the drugs in which risks will eventually overshadow their benefits, although past experience suggests that it is unlikely. Dr Lexchin should be requested to substantiate his apparent inference that the greater use of safety warnings indicates approval of unsafe drugs.

References

1. Appel WC. Adverse drug reaction reporting controversy. *CMAJ* 2002; 166: 884-5.
2. Rawson NSB, Ross Terres JA. Rosiglitazone use and associated adverse event rates in Canada between 2004 and 2010. *BMC Res Notes* 2013; 6: 82.
3. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA* 1998; 279: 1200-5.
4. Fremont-Smith K. Adverse drug reactions in hospitalized patients. *JAMA* 1998; 280: 1741.
5. Kravitz GR. Adverse drug reactions in hospitalized patients. *JAMA* 1998; 280: 1741.
6. Kvasz M, Allen IE, Gordon MJ, et al. Adverse drug reactions in hospitalized patients: a critique of a meta-analysis. *Med Gen Med* 2000; 2: E3.
7. Rawson NSB. Adverse drug reactions in Canada: facts *v* urban myths. *Canadian Health Policy*. Toronto: Canadian Health Policy Institute, October 2013.
8. Lexchin J. New drugs and safety: what happened to new active substances approved in Canada between 1995 and 2010? *Arch*

	Intern Med 2012; 172: 1680-1. 9. Rawson NSB. New drug approval times and safety warnings in the United States and Canada, 1992-2011. *J Popul Ther Clin Pharmacol* 2013; 20: e67-81. 10. Rawson NSB. Does the increase in safety warnings from regulators *actually* mean that more dangerous drugs are being approved? *Canadian Health Policy*. Toronto: Canadian Health Policy Institute, August 2013. 11. Lessons from cisapride. *CMAJ* 2001; 164:1269.

REVIEWER	Robin Ferner West Midlands Centre for Adverse Drug Reactions Birmingham England
REVIEW RETURNED	18-Nov-2013

GENERAL COMMENTS	The author asks an interesting question regarding the safety of medicines, which he divides into three groups: traditional, significant therapeutic advances, and biologics. The division is based on the grading by the Canadian Patented Medicine Prices Review Board and, failing that, Prescrire International (which can sometimes be tendentious). If I have understood correctly, the extent to which these organizations concurred is unclear, since Prescrire was only examined if PMPRB had provided no grading. There is therefore some uncertainty about the initial classification. The major difficulty in reaching a conclusion is that warnings and withdrawals are used as a measure of safety. (1) Warnings often refer to remote or rare harms (oligohydramnios with transtuzumab, for example). (2) The harms already noted in the product information at launch will not subsequently be added as warnings to the label. (3) The aggregate harm will depend on the number and seriousness of injurious reactions occurring in treated patients, not on the number of warnings. The K-M curves, especially Figure 2, suggest that new active substances representing a significant advance may well be more likely to generate warnings, but the study is underpowered. Since there is no power calculation, that view cannot be examined. Of course, the data depend on the number of drugs available for study, so it is not a criticism of the work that the study is underpowered; but it is a criticism that it may give false reassurance. The question is an interesting one, and with suitable statistical and descriptive cautions about the gap between what has been examined and the ideal, it could be published

VERSION 1 – AUTHOR RESPONSE

Thomas Moore

A statistical study based on “proving” rather than disproving the null hypothesis succumbs to Type II statistical error unless accompanied by a substantial analysis capable of identifying the characteristics of the differences, should they exist. Otherwise it is impossible to tell whether the study truly concludes there is no difference, or that the methodology was sufficiently flawed that it could not detect a true difference. In drug bioequivalence assays and non-inferiority clinical trials, for example, a finding of equivalence or non-inferiority is supported by such analysis, although even these methods are not without controversy. These studies also contain validated endpoints.

I thank Mr. Moore for pointing out that I should have been talking about disproving the null hypothesis and that error has been corrected. However, the point about a Type II error only applies if I was taking a sample whereas in both of the comparisons I am looking at the entire population. I have made it clear in the Methods section that I am dealing with an entire population.

The next issue concerns the conclusions reached that there is “no difference in safety” of biological products compared to small molecule drugs. (p13/line 25). Next at line 32 the finding is described as “postmarket safety,” a different concept. In the next paragraph it is further expanded to say “the safety profile of NAS is the same regardless of the level of therapeutic benefit.”

Part of this is a language problem. The investigator has observed post-market safety warning actions by a regulator. That is not a proxy for “safety” or a “safety profile,” and I suggest they should not be considered synonyms. The safety profile of a drug needs to include any boxed warnings, contraindications and other risks detected prior to approval as well as post approval.

Again I’d like to thank Mr. Moore for pointing out the inconsistency in wording and this has been corrected. I also agree with him that the overall safety of a drug includes both problems identified before approval and after approval, but this is a study of postmarket safety. In the Discussion I note the difference between postmarket safety warnings and the overall safety of a drug.

The third limitation with the safety warning metric is the wide variety in the incidence and severity of safety problems involved. A warning might indicate a handful of reported cases of a catastrophic side effect (efalizumab, progressive multifocal leukoencephalopathy), a problem for a small patient subset (fluoroquinolones, worsening of myasthenia gravis), or a massive contraindication (dolasetron – any therapeutic use under 18 years age, any post operative nausea).

Mr. Moore is correct that not all safety warnings are equivalent but trying to rank them would introduce some very subjective decision making, e.g., is a very severe safety issue affecting a small number of people worse than a less severe safety issue affecting a large number of people. The point about all safety warnings being treated as equivalent is now listed as a limitation of the paper. Moreover, the

methodology that I use – either the withdrawal of the drug or a first serious safety warning, regardless of the nature of the safety problem, has been used by a number of other researchers, see: Lasser et al. JAMA 2002;287:2215-2220; Giezen et al. JAMA 2008;300:1887-1896; Arnardottir et al. British Journal of Clinical Pharmacology 2011;72:490-9. In the Discussion I note that the methodology that I used is consistent with the methodology used in the papers by Lasser, Giezen and Arnadottir.

The comparison involving therapeutic advances also appear to be definition sensitive. In the Archives study the author used a priority review as the primary index of therapeutic value and reported a substantial and statistically significant difference between priority and standard review drugs and subsequent safety warnings.

In the current study, a different definition of therapeutic value was created using two versions of a Canadian formulary review board criteria for some agents and recode of ratings from the French journal Prescrire for others. In the new definition the differences seen in the previous study do not reach statistical significance.

I'm somewhat puzzled by the therapeutic value issue. Given a subjective judgment, it does not seem surprising to discover differences, particularly with different criteria for designating a novel drug. And using 3 different rating schemes for one variable does not help, and is probably not the best solution to addressing the missing value problem.

Mr. Moore seems to assume that the decision to review a new drug submission through the priority pathway is an objective decision but while there may be some factual basis behind the decision there is also a subjective component. I agree that ratings of therapeutic value also carry a subjective element but I don't think that the subjectivity in this decision is any greater than it is when the regulator decides to put the drug in the priority track.

The Patented Medicine Prices Review Board is not a formulary committee. It is a body that sets the maximum introductory price for a new patented drug and as part of that process it evaluates the therapeutic worth of these products. It is true that it changed its rating scale in 2010 but that change, which is described in the manuscript, primarily created more gradations in the not innovative group of drugs and so did not affect the dichotomous rating of innovative or not innovative group. The point about the change in the PMPRB rating scale not affecting the difference between innovative and not innovative is made in the Methods. Prescrire ratings were used to increase the sample of drugs that could be classified as innovative or non-innovative. Previous work has found a moderate degree of agreement between PMPRB and Prescrire evaluations.

Also, some distortion is introduced by using the median time to first safety warning, rather than the median and range for all the serious safety warnings. The latter seems to be a more accurate index of when prescribing physicians will learn about drug risks not detected or fully understood at time of approval. This complicates a survival analysis, but measuring all warnings probably provides a more accurate measure of how long after approval it takes to detect substantial new safety issues. In two published studies in the US the median time to all warnings was 11 years (compared to 4 years to first

warning in these data).

The decision to use time to first safety warning is in line with the methodology used in three other studies referred to above. This point is made in the Discussion.

Nigel Rawson

Dr Lexchin begins his article by stating that drug safety is becoming a topic of increasing concern in Canada. However, since the thalidomide tragedy, drug safety has been of ongoing concern to Canadians, which increases periodically when drugs, such as practolol, benoxaprofen, troglitazone, cisapride, cerivastatin and rofecoxib, taken by large numbers of patients are withdrawn for safety reasons. What is of greater concern is how Health Canada identifies drug safety issues, with only an under-funded voluntary adverse drug reaction (ADR) reporting system that can in no way be regarded as a reliable safety net, and the frequent lack of clarity and timeliness in its communications about drug risks.

Dr. Rawson mentions drugs that were never approved in Canada (practolol) or were approved but never marketed (troglitazone). He is correct that Health Canada is grossly underfunded to detect ADRs but ADR detection is only one of the tools used by Health Canada in its decision to either issue a serious safety warning or to withdraw a drug from the market. This point is made in the Discussion. Dr. Rawson is correct that concern about safety issues flares up from time to time but the literature that I cite indicates that concern about this issue has been more sustained over the recent past and therefore I believe that my statement is accurate. Since the original draft of this paper was prepared the federal government has introduced legislation giving it the power to order postmarket safety studies and I have referred to this legislation in the Introduction.

Dr Lexchin also states that ADRs "are the fourth to sixth leading annual cause of death." This assertion is based on a 1998 meta-analysis that was criticized at the time of its publication as significantly over-estimating the number of fatal ADRs due to using studies too diverse to aggregate. Moreover, 60% of the studies were published before 1980 and had higher incidences of ADRs, whereas the later studies had much lower incidences. A repeat meta-analysis using the same methods performed in 2000 by experts in the technique demonstrated that ADRs were not defined with sufficient precision and consistency across the studies to permit the pooling of results, correct numbers were not used for calculating their incidence rates, and appropriate adjustments were not made for differences in studies, patients and outcomes, which likely led to over-estimation of the number of patients affected. It is time for this urban myth to be abandoned.

The number of deaths annually in United States hospitals due to ADRs that is found in the Lazarou study is quite similar to the figure that is reported in a 2012 publication by the Institute of Safe Medication Practices (ISMP, "QuarterWatch", Moore et al, 2012). The reference has been added in the Introduction. The repeat meta-analysis that Dr. Rawson references was published in a journal that was only published irregularly between 1999 to 2007. It is not clear that the journal was peer reviewed and the study received financial support from the Pharmaceutical Research and Manufacturers of America, an organization that creates a clear conflict-of-interest since it is in PhRMA's interest to downplay the number of deaths due to ADRs caused by its products. I have not mentioned the Kvasv paper or critiqued it in the revision.

Dr Lexchin uses data on 406 new drugs approved in Canada between January 1997 and March 2012 obtained from Health Canada's annual performance reports and links them with information on serious safety warnings from the agency's MedEffect website "supplemented by searching on the product name in the Drug Product Database (DPD)". The DPD allows access to the relevant Product Monographs (PMs), which are the definitive legal documents on drugs in Canada and include all warnings (serious or not). Serious warnings are usually highlighted by a box but, in older PMs, bold capital letters alone may be used. Dr Lexchin groups the drugs using the Patented Medicine Prices Review Board classification of new drugs, which is an exceedingly restrictive categorization such that few products are considered to be breakthroughs or substantial

improvements. If a drug for the disease or condition exists, whether it is highly or poorly effective, it is rare for a new, more effective drug to be considered as a breakthrough or substantial improvement. Dr Lexchin assesses, as his principal objectives, the survival rates of the drugs until they acquire a serious safety warning or are discontinued for a safety reason, comparing biologic versus traditional drugs (although the performance reports "only gave the type of product … from 2000 onwards") and drugs with "significant therapeutic advances" versus those without.

Dr. Rawson takes issue with the use of the Patented Medicine Prices Review Board characterization of the therapeutic value of new drugs. Overall the PMPRB finds that about 1 in 10 new patented drugs have significant therapeutic value compared to existing products. That 10% figure is consistent with the figure from Prescrire (see: New drugs and indications in 2010: inadequate assessment; patients at risk. Prescrire International 2011;20:105-110) and with an assessment of drugs for general practice done by a Belgian drug bulletin (personal communication cited in Lexchin J. Health Policy 2012;105:221-225). More importantly, the criteria that the PMPRB uses have nothing to do with whether a drug receives a postmarket safety warning. Therefore, I have not addressed the issue of the PMPRB classification system in the revision.

Dr Lexchin finds that, of 406 new drugs approved between January 1997 and March 2012, 87 (21.4%) acquired a serious safety warning or were discontinued for a safety reason. This is similar to the rate of 19.4% that he reported elsewhere when he examined, using the same methods, 434 new drugs approved in Canada between January 1995 and December 2010.

However, these figures are differ substantially from those I recently reported based on a study of the approval times and the acquisition of a serious safety warning or discontinuation for a safety reason for 454 new therapeutic (i.e. new salts, esters, dosage forms and combinations of previously approved drugs, diagnostic products and vaccines were excluded – it is unclear whether Lexchin had any exclusions) approved between 1992 and 2011 in both Canada and the United States.

The manuscript clearly outlines that only drugs considered as New Active Substances by Health Canada were included.

Drugs discontinued for safety reasons were identified from websites of Health Canada and US Food and Drug Administration (FDA) and publications. Serious safety warnings in Canada were identified from the relevant PM obtained from the DPD website, supplemented with information from the MedEffect website. Black-box warnings (BBWs) were obtained from the history of each drug's Label available on the FDA's website, supplemented with information from the FDA's MedWatch website. In contrast to Dr Lexchin's results, I found that 199 (43.8%) of the new drugs acquired a serious safety warning or were discontinued for a safety reason in Canada, while 158 (34.8%) acquired a BBW or were discontinued for a safety reason in the United States.

This difference in the rate of acquisition of serious safety warnings leads to substantial differences in the probability of acquiring a serious safety warning or safety discontinuation (Table 1) between biologics and traditional drugs. Lexchin found no statistical difference between the two types of drugs (log-rank test, $p=0.47$), whereas I found significant differences between them in both Canada and the United States (both log-rank tests, $p<0.0001$).

Why the difference in results? I believe that the use of PMs, supplemented by MedEffect information, is a more comprehensive assessment of the acquisition of serious safety warnings than the use of MedEffect alone. MedEffect frequently omits serious drug warnings. The Canadian data using PMs (plus MedEffect) are also consistent with US data using Labels (plus MedWatch), although the Canadian rates are higher. Dr Lexchin's article is deficient without some discussion of the difference between his methods and results and mine and the limitations imposed by relying mainly on MedEffect.

The crux of these comments is that Dr. Rawson and I reached different conclusions about the rate of acquisition of serious safety warnings for drugs in Canada and that is correct. Dr. Rawson is also correct that the reason for this difference is that we used different methodologies. He combined safety warnings from Canada's MedEffect web site with bolded safety warnings in the Product

Monograph. However, he only looked at the most recent version of the Product Monograph – he did not have the Product Monograph when the product was first marketed and therefore he had no way of knowing whether the safety warnings in the Product Monograph were identified in the premarket approval process or in the postmarket phase. Therefore, he did not do a study on just postmarket safety warnings. Because of the significant difference in the design of Dr. Rawson's study and mine I have not mentioned his study in the revision.

The title of Dr Lexchin's article seems to imply that an increase in the use of serious safety warnings would indicate that unsafe drugs are being approved. However, safety warnings are exactly that – *warnings*. They are not indicative of actual health outcomes. The only solid evidence about overall drug safety is the rate of discontinuation for safety reasons, which in drugs approved in 2002-2011 was 2.1% in Canada and 1.3% in the United States, which was marginally less than the corresponding rates of 2.3% and 1.7%, respectively, in 1992-2001.

I fail to see how Dr. Rawson could take the title of the article and conclude from it that I was claiming that Health Canada approves unsafe drugs. However, he is correct safety warnings do not tell us about the overall safety of drug. As I said in response to Mr. Moore that point is now made.

Drugs are not approved by regulators unless their benefits outweigh their risks. Health Canada and the FDA have been working to improve the timeliness of the approval of innovative drugs in order to get them as quickly as is safely possible to patients. One way that they have done this is by increasing the number of drugs receiving a priority status review. Many of these products are innovative pharmacologically-complex drugs that may provide novel benefits but may also have unusual risks. To ensure that patients and healthcare providers are aware of this, Health Canada and the FDA have increasingly required serious warnings in the PMs and Labels over the last decade, frequently at the time of approval. An increase in their use does not necessarily mean that unsafe drugs have been approved.

This statement has nothing to do with my manuscript and doesn't require a revision in the paper.

Robin Ferner

The division is based on the grading by the Canadian Patented Medicine Prices Review Board and, failing that, Prescrire International (which can sometimes be tendentious).

If I have understood correctly, the extent to which these organizations concurred is unclear, since Prescrire was only examined if PMPRB had provided no grading.

There is therefore some uncertainty about the initial classification.

Dr. Ferner is correct that the decision about whether a product is innovative is based on the ratings from the PMPRB and when it had not examined the drug then the rating from Prescrire International was used. I am not sure what he means when he refers to Prescrire International as sometimes tendentious.

In previous work I have examined the agreement between the PMPRB and Prescrire ratings and I cite that work in the manuscript.

The major difficulty in reaching a conclusion is that warnings and withdrawals are used as a measure of safety.

(1) Warnings often refer to remote or rare harms (oligohydramnios with transtuzumab, for example).

(2) The harms already noted in the product information at launch will not subsequently be added as warnings to the label.

(3) The aggregate harm will depend on the number and seriousness of injurious reactions occurring in treated patients, not on the number of warnings.

Like the other two reviewers Dr. Ferner is correct that it would be far better to have data on the number of people harmed rather than to just look at safety warnings but that information is not available. That point is made when discussing limitations. Furthermore, as I pointed out in my response to Mr. Moore, the methodology that I used has been used in three other published studies.

The K-M curves, especially Figure 2, suggest that new active substances representing a significant advance may well be more likely to generate warnings, but the study is underpowered. Since there is no power calculation, that view cannot be examined. Of course, the data depend on the number of drugs available for study, so it is not a criticism of the work that the study is underpowered; but it is a criticism that it may give false reassurance.

There is no power calculation because I am not using a sample but an entire population. Also a power calculation assumes that there is some basis for deciding a priori what the difference might be and since this is a novel study there was no prior knowledge on this topic.

The question is an interesting one, and with suitable statistical and descriptive cautions about the gap between what has been examined and the ideal, it could be published.

I am unclear as to what statistical and descriptive cautions Dr. Ferner is referring to.

VERSION 2 – REVIEW

REVIEWER	Nigel Rawson Eastlake Research Group Canada I have published results of a study assessing the acquisition of serious safety warnings in Canada and the United States that has higher rates than in Dr Lexchin's study that are not considered.
REVIEW RETURNED	20-Jan-2014

GENERAL COMMENTS	I am unable to say that the references are up-to-date or that the study limitations are discussed adequately without Dr Lexchin providing comments regarding the differences between my work and his in his article. Dr Lexchin responds that the repeat meta-analysis received financial

	support from the pharmaceutical industry association, which may be an issue. However, it was performed by experts in meta-analysis. The crux of the matter comes down to whether one believes in work done by an MSc student and his supervisors that was criticized at the time of publication as exaggerating the rate of deaths in hospital from adverse drug reactions, or work done by experts in meta-analysis that received industry support. With regard to the differences between Dr Lexchin's estimates of the rate of acquisition of serious safety warnings for drugs in Canada and mine,1 Dr Lexchin responds that the methodologies were different (in which he is correct) and that I only looked at the current Product Monograph (PM) because Health Canada does not make earlier versions available on its website. Information in the latest PM was adequate for all but 48 of the 454 drugs in my study for which I had to make an assumption. However, Dr Lexchin may not have noted that I was able to obtain all PMs for 25% of the 48 drugs (workload at Health Canada prevented access to all PMs for all 16 drugs) and my assumption was valid for all but 1 of the 16 drugs. In my study, I included all serious warnings in the PM and in MedEffect. I, therefore, do not accept his statement that I "did not do a study of just postmarket safety warnings." References 1. Rawson NSB. New drug approval times and safety warnings in the United States and Canada, 1992-2011. J Popul Ther Clin Pharmacol 2013; 20: e67-81.

REVIEWER	Robin Ferner West Midlands Centre for Adverse Drug Reactions
REVIEW RETURNED	20-Jan-2014

GENERAL COMMENTS	I recommended publication of the original version with some 'statistical and descriptive cautions.' The author asks for clarification. These are changes required in the current version: A. The conclusion detailed in the abstract should be altered from 'There is no difference in postmarket regulatory safety action between traditional medications and biologics and no difference between drugs with significant therapeutic benefits and those without' to 'There *was* no difference postmarket regulatory safety action between traditional medications and biologics and no difference between drugs with significant therapeutic benefits and those without.' since there is insufficient evidence that the results can be generalized. B. The author should note that results do not exclude striking differences between groups. For example, % warning/withdrawn traditional:biologic could plausibly be anywhere from 40.2/18 to 21.8/39.7 according to the confidence intervals, that is from 2x to 1/2x. This is another way of saying that the estimates of % are not precise.

	C. The difficulties of linking regulatory change with real harm, discussed by the author in his responses to all three referees, are not discussed under 'limitations,' but should be.

VERSION 2 – AUTHOR RESPONSE

Response to Dr. Ferner

A. The conclusion detailed in the abstract should be altered from

'There is no difference in postmarket regulatory safety action between traditional medications and biologics and no difference between drugs with significant therapeutic benefits and those without'

to

'There *was* no difference postmarket regulatory safety action between traditional medications and biologics and no difference between drugs with significant therapeutic benefits and those without.'

since there is insufficient evidence that the results can be generalized.

This change was made.

B. The author should note that results do not exclude striking differences between groups. For example, % warning/withdrawn traditional:biologic could plausibly be anywhere from 40.2/18 to 21.8/39.7 according to the confidence intervals, that is from 2x to 1/2x. This is another way of saying that the estimates of % are not precise.

The following qualification was inserted into the first paragraph of the Discussion: "However, given the wide confidence intervals this conclusion should be regarded as tentative. Further research using a larger number of NAS might show differences."

C. The difficulties of linking regulatory change with real harm, discussed by the author in his responses to all three referees, are not discussed under 'limitations,' but should be.

The following sentence has been added to the end of the paragraph on limitations: "Finally, it is important to note that the regulatory decision to issue a safety warning should not be equated with the actual degree of harm caused by the drug."

Response to Dr. Rawson

Dr Lexchin responds that the repeat meta-analysis received financial support from the pharmaceutical

industry association, which may be an issue. However, it was performed by experts in meta-analysis. The crux of the matter comes down to whether one believes in work done by an MSc student and his supervisors that was criticized at the time of publication as exaggerating the rate of deaths in hospital from adverse drug reactions, or work done by experts in meta-analysis that received industry support.

The citation of the Lazarou et al study is a relatively minor point in my article but one that Dr. Rawson seems particularly concerned about. To begin with whether or not Lazarou was an MSc student when the article was published is not particularly relevant. Strong work has been published by MSc students (see Roseman et al. JAMA 2011;305:1008-17; Roseman et al. BMJ 2012;345:e5155 – for full disclosure I was one of the co-authors on these papers). When the Lazarou article was published it was accompanied by an editorial. The editorial did challenge some of the conclusions but also said "the authors adhered to generally accepted criteria for meta-analyses" and "these data are important, and even if the true incidence of ADRs is somewhat lower than that reported by Lazarou et al, it is still high". There were four subsequent letters criticizing the article but also one supporting it and the criticisms were addressed by Lazarou and his co-authors. Contrary to what Dr. Rawson says, the issue is not only about who did the meta-analyses but also about the credibility of where they were published. Lazarou's was published in a highly respected high impact peer reviewed journal whereas the one by Kvasv et al came out in a journal that published irregularly, is no longer being published and that may or may not have been peer reviewed. Therefore, I have not made any changes to the citation of the Lazarou et al article.

With regard to the differences between Dr Lexchin's estimates of the rate of acquisition of serious safety warnings for drugs in Canada and mine,1 Dr Lexchin responds that the methodologies were different (in which he is correct) and that I only looked at the current Product Monograph (PM) because Health Canada does not make earlier versions available on its website. Information in the latest PM was adequate for all but 48 of the 454 drugs in my study for which I had to make an assumption. However, Dr Lexchin may not have noted that I was able to obtain all PMs for 25% of the 48 drugs (workload at Health Canada prevented access to all PMs for all 16 drugs) and my assumption was valid for all but 1 of the 16 drugs. In my study, I included all serious warnings in the PM and in MedEffect. I, therefore, do not accept his statement that I "did not do a study of just postmarket safety warnings."

Even after a close reading of Dr. Rawson's article it is difficult to fully understand his methodology. He says that if there was a black box warning in the Canadian product monograph but this warning did not appear in the MedEffect database and there was also such a warning in the American product label then he assumed that the earliest appearance of the American warning was also the time when the Canadian warning appeared. He goes on to say that he tested the validity of this assumption by requesting the original product monographs for 25% of the 48 products where he made the assumption and that four of those drugs had a black box warning at the time of approval. It is unclear if those four were included in his analysis; if they were then these were not safety issues identified postmarket. I also question his assumption that Health Canada would add a serious safety warning to the Product Monograph without issuing a warning to the public and physicians through the MedEffect web site. (Dr. Rawson could have identified black box warnings that were present when the product was approved by looking at the first listing of the product in the Compendium of Pharmaceuticals and Specialties.) Rawson's article identifies bicalutamide as being withdrawn from the Canadian market but in fact it was only the Notice of Compliance with conditions for a particular indication that was withdrawn, the drug is still available on the Canadian market. Besides bicalutamide he lists 14 other drugs withdrawn in Canada between 1992 and 2011 but he misses ceftobiprole, dexfenfluramine,

gatifloxacin, lumiracoxib, remoxipride and sitaxsentam. The misidentification of bicalutamide and the missing products raise questions about the thoroughness of his search. A debate between Dr. Rawson and myself might prove to be very interesting but I do not believe that my article is the appropriate place for that debate and therefore I have not included a description of his article or my critique of it in this revision.