PERSPECTIVES

Should benefit–risk assessment have its own drug "label"?

R Scott Braithwaite

Section of Value and Comparative
Effectiveness, Division of General
Internal Medicine, New York
University School of Medicine,
New York, NY, USA

Abstract: Many consumers and clinicians incorrectly believe that the Food and Drug Administration (FDA) approval of a new therapeutic implies that its benefits have been proven to exceed its harms. While the FDA could require proof that benefits exceed harms prior to approval, it has been argued that this approach would be infeasible because of prohibitively large sample sizes. One possible alternative would be for the FDA to supplement its standard "label" denoting "safe and effective" with a secondary "label" denoting benefits have been demonstrated to exceed harms, which would be granted only after sufficient post-marketing data had accumulated to prove that its benefits exceeded its harms. This secondary label would not necessarily be linked to marketing restrictions or other commercial prohibitions but, rather, would be only information for consumers and clinicians. Strengths, weaknesses, and feasibility challenges of this approach are discussed.

Keywords: drug label, Food and Drug Administration, safety, efficacy, benefit–risk assessment

Introduction

Many consumers may believe that the Food and Drug Administration (FDA) approval of a new therapeutic implies that its benefits have been proven to exceed its harms. Indeed, the FDA's "For Consumers" website encourages precisely that interpretation, stating "[i]f FDA grants an approval, it means the agency has determined that the benefits of the product outweigh the risks for the intended use."[1] Arguably, much of the backlash directed at the FDA after the withdrawal of over 12 high-profile brand-name drugs (eg, rofecoxib, troglitazone)[2–5] originated because consumers interpreted these withdrawals as proof that a presumptive judgment about benefits exceeding harms was inaccurate and premature.[6] However, it is well documented that FDA approval does not indicate proof that benefits exceed harms.[7] A marginally effective but "safe" drug may still have harms that exceed its benefits and, conversely, an unequivocally effective but "unsafe" drug may have benefits that exceed its harms. Magnifying confusion about these two distinct ideas, the FDA considers available information about benefit–risk balance at the same time that it assesses safety and efficacy for approval decisions, although risk/benefit information is applied implicitly and unsystematically rather than explicitly, systematically, and quantitatively.[8,9] Consequently, the FDA does not require "proof" that benefits exceed harms, whereas it does require proof of safety and effectiveness.

While it can be argued that the FDA should make proof that benefits exceed harms a prerequisite for approval, this criterion may be infeasible because detecting unknown,

Correspondence: R Scott Braithwaite
Section of Value and Comparative
Effectiveness, Division of General Internal
Medicine, New York University School of
Medicine, 423 E 23rd Street 15th Floor,
New York, NY 10010, USA
Tel +1 212 263 4964
Fax +1 212 263 4983
Email scott.braithwaite@nyumc.org

infrequent events sometimes requires huge sample sizes that are impractical for pre-marketing studies. However, an alternative approach to make benefit/harm assessment more explicit could be to create a new label that would be conferred only after sufficient evidence has accumulated to prove that benefits exceed harms.

A new label for "benefits exceed harms"?

The FDA could supplement its standard label denoting safety and efficacy with a secondary label denoting benefits exceeding harms. Unlike the "safe and effective" label, this secondary "benefits exceed harms" label would not necessarily be linked to marketing restrictions or other commercial prohibitions but, rather, would be only information for consumers and clinicians, and could be harmonized with other approaches, such as the "drug box," to make drug information more transparent to consumers.[10] Because this secondary label would evaluate a concept with immediate relevance for decision making rather than the more abstract and context-dependent idea of "safety,"[8] it could lead to more informed shared decision making for consumers and physicians.

Determining whether benefits exceed harms

There are many systematic frameworks for quantitatively estimating whether benefits exceed harms[6–9] (benefit–risk assessment [BRA]). While the purpose of this paper is not to systematically review BRA frameworks or to advocate using one framework over another, I will choose one particular BRA approach, Incremental Net Health Benefit (INHB), for illustration.[7] This approach quantitatively compares benefits and harms on the same scale, and explicitly considers the idea that regulators or consumers may be risk averse (ie, weigh a harm more than they would weigh a benefit of equal magnitude). (Briefly, the INHB of Therapy 2 versus Therapy 1 can be expressed as $INHB = (E2 - E1) - (R2 - R1)$, where effectiveness [E] and risk [R] are measured in the same units. A "favorable" benefit–risk balance occurs when $(E2 - E1) > (R2 - R1) + X$ or $(E2 - E1) > (R2 - R1) * X$, where "X" is an additive or multiplicative factor that can reflect the risk aversion of regulators and can increase the mandated margin by which benefits exceed risks.)[7]

Example

Suppose the FDA grants pre-marketing approval for a weight-loss drug based on pre-marketing studies that show clinically significant improvements in 20% of patients.

Further, assume that the 20% improvement is a reasonable approximation of the true biological effect. The pre-marketing studies did not suggest any serious adverse events, and the drug's biological mechanism did not raise any concerns about particular harms. However, assume that the drug confers a true biological harm, life-threatening bleeding, in 0.2% of patients. This adverse event was too rare to be detected in the two pre-marketing studies, each of which enrolled 800 patients.

What data would be sufficient to determine that this drug has a favorable BRA? Regulators could use INHB to address this question by asking how many people would need to have the demonstrated incremental improvement in weight loss to offset one person having a serious adverse event from the drug. While one of the challenges of applying BRA to a new therapeutic is that the adverse event profile of a new therapeutic may be unknown, it is possible to test a plausible range of assumptions regarding the morbidity and mortality profile of hypothetical unknown adverse events, and to assess the sample size and/or follow-up necessary to have adequate statistical power to detect them.

For example, applying INHB with substantial risk aversion might suggest that ≥ 100 persons would need to have clinically significant weight loss to offset one person having serious harm. Consequently, to evaluate the suitability of a "benefits exceed harms" label, regulators would need to determine whether sufficient data have accumulated to rule out an adverse event occurring at a corresponding frequency (eg, in this case, at least 1/100 of the frequency of the anticipated benefit, or two events per 1000 patients). Using a standard power criterion of 80% and assuming a baseline frequency of one event per 1000 patients, this threshold would be attained after 12,340 patients have been studied. Therefore, this weight-loss drug could go on the market with FDA approval after the standard pre-marketing studies are concluded but would not receive a "benefits exceed harms" label until data have accumulated on at least 12,340 patients and until these data are investigated for unanticipated harm signals (Table 1).

It is important to note that while the "safe and effective" label would generally precede the "benefits exceed harms" label, this is not necessarily the case. Indeed, for a new therapeutic that confers a very large incremental benefit, the number of observations needed to demonstrate that benefits exceed harms would be small, and the "benefits exceed harms" label could be granted before the "safe and effective" label is granted (in this context, it could fill the role of the current "experimental use" designation).

Table 1 Suppose evidence is being evaluated regarding a weight loss drug that confers a clinically significant benefit in 20% of patients treated but yields a life-threatening harm in 0.2% of patients treated. Using Incremental Net Health Benefit to infer that one patient harmed would offset at least 100 patients benefitted and assuming that the harm has a baseline frequency of 0.1%, any study with an n between 186 and 12,340 would have sufficient power to detect the benefit, yet would have insufficient power to detect the harm. Consequently, a label designating proof that "benefits exceed harms" could only be justified based on data well beyond those demonstrating benefits or supporting regulatory approval

n (both arms)	FDA approval?	Sufficient power to detect benefit?	Sufficient power to detect harms?	Sufficient power to detect harms offsetting benefit?
n < 186	No	No	No	No
186 ≤ n < 1800	No	Yes	No	No
1800 ≤ n < 12,340	Yes	Yes	No	No
n > 12,340	Yes	Yes	Yes	Yes

Abbreviation: n, number.

Challenges to a new label for "benefits exceed harms"

BRA approaches in general, and the INHB method in particular, require embedded assumptions about what level of risk aversion is most suitable for a regulatory authority, and about what level of statistical certitude should be required.[7] For example, it may be argued that the standard criterion for null hypothesis testing in clinical studies (requiring P values < 5%) is too strict for disproving the analogous null hypothesis for BRA (ie, that benefits do not exceed harms), and therefore alternative criteria should be preferred.

It might be suggested that societal or regulatory preferences regarding BRA may differ greatly from individual patient preferences and risk tolerances and, therefore, it is neither possible nor desirable for regulatory authorities to attempt to make a particular BRA decision on behalf of society.[11] However, it is important to appreciate that analogous arguments may be made for most other regulatory activities. Individuals may have greatly differing definitions and evidence standards for "safety" and "effectiveness," for example, yet the FDA implicitly chooses a particular societal standard when it approves therapeutics. Indeed, one way to mitigate this concern is to make explicit in a "benefits exceeding harms" label that individual preferences, risks, and benefits vary; therefore, patients should make an individualized and shared decision with their clinician (Figure 1). For example, risk-tolerant patients may choose to try a new effective drug even before its benefits have been proven to exceed its harms, because they may be satisfied with a reasonable likelihood, rather than proof, that benefits exceed harms, particularly if they have a high symptom burden.

It may be appreciated that evaluations of incremental benefits and harms will not be static, even without new evidence generation, because comparators may change over time as new alternatives become available and old alternatives are removed from the market. However, secondary labels can be revisited at suitable time intervals (eg, 5 years) that are used in other situations in which evidence reviews need to be updated, such as systematic reviews used for updating clinical guidelines.

Some may argue that backlash about withdrawn drugs has occurred because of unknown effects, and such effects would not have been included in a quantitative

A

> This therapy has been proven to be effective and relatively safe. However, it has not been studied long enough to prove that its benefits exceed its harms for a typical person that this therapy is approved for. Discuss your personalized risk of benefits and harms with your clinician.

B

> This therapy has been studied long enough to prove that its benefits exceed its harms for a typical person that this therapy is approved for. However, every patient is different. Discuss your personalized risk of benefits and harms with your clinician.

Figure 1 Language that could be included as part of a post-marketing "benefits exceed harms" label. Prior to the issuing of this label ([**A**] possible lay explanation), data may have been consistent with benefits exceeding harms, but not robust enough to constitute proof. Issuing the "benefits exceed harms" label ([**B**] possible lay explanation) would designate that sufficient evidence had accumulated to allow detection of harms that could offset known benefits and, consequently, proof that benefits exceed harms. This is a far higher evidence standard for benefit–risk assessment than is currently required for marketing approval, and is more akin to the current evidence standard for effectiveness.

benefit/harm analysis. However, this argument does not consider the idea that a quantitative benefit/harm analysis should only be performed after accumulation of evidence with sufficient statistical power to detect harms that would outweigh observed benefits. Backlash about unknown effects has generally occurred when drugs were approved before evidence accumulated with sufficient statistical power to detect those rare harmful events.

Finally, it may be questioned whether adding a second "benefits exceed harms" label would be preferable to classifying a newly marketed drug as "experimental" until sufficient post-marketing data have accumulated to prove that benefits exceed harms. Indeed, expanding the use of the "experimental" designation to all situations in which benefits have not yet been proven to exceed harms may accomplish the same goal. However, expanding the experimental designation in this manner would likely result in many new drugs coming to market labeled "experimental." Such a sudden and sharp change in the use of a term currently associated with terminal illness could lead to substantial confusion. In contrast, adding a "benefits exceed harms" label could avoid this confusion and, indeed, may clarify the distinction between the separate criteria of "safe and effective" and "benefits exceed harms."

Advantages of a new label for "benefits exceed harms"

Developing a second label for "benefits exceed harms" may lead to more informed decision making about "off-label" uses for therapeutics. For example, a therapeutic approved for Condition A may receive a "benefits exceed harms" label for Condition A, but still be used off-label for Condition B. While this off-label use may continue to be permitted, consumers and clinicians would potentially be more aware that benefits have not been demonstrated to exceed harms for these off-label uses. Indeed, this growing awareness may lead manufacturers to seek "benefits exceed harms" labels for uses that had primarily been off-label and may encourage more circumspect utilization of new therapeutics beyond their approved indication.

It is possible that a "benefits exceeding harms" label would incentivize longer-term adverse-event monitoring and reporting by manufacturers of new therapeutics, as this longer-term follow-up would often be necessary to generate the additional evidence to establish that benefits exceed harms. Indeed, if the FDA adopted a "benefits exceeding harms" label, this may facilitate a more systematic assessment of post-marketing data to look for harm signals.

Other considerations affecting a new label for "benefits exceed harms"

It is important to observe that absence of a "benefits exceed harms" label for a particular therapeutic need not preclude payers' reimbursement for that therapeutic. Even when benefits have not been demonstrated to exceed harms overall, benefits might be likely to exceed harms for a particular patient in a particular situation. However, payers may wish to require evidence of a shared decision between consumer and clinician that considers known benefits and harms (analogous to informed consent) before reimbursing therapeutics that do not have the "benefits exceeds harms" label.

Finally, it can be appreciated that circumstances in which additional evidence would be necessary to establish that benefits exceed harms are also circumstances where a formal analysis of expected value of information would support additional gathering of evidence, unless both benefits and risks are very small. In these circumstances, additional information would provide clearer inferences for decision making. Conversely, situations where benefits have been shown to exceed harms are situations in which the value of information will not support additional research because there is already a clear inference for decision making.[11]

Conclusion

If it is infeasible for the FDA to require proof that benefits exceed harms prior to marketing approval, it is worth asking whether the FDA should supplement its standard label denoting "safe and effective" with a secondary label denoting "benefits have been demonstrated to exceed harms." Because this secondary label would evaluate a concept with more relevance for decision making than the abstract and context-dependent idea of "safety," it could lead to more informed shared decision making for consumers and physicians.

Disclosure

References

1. US Food and Drug Administration. Is it really FDA Approved? February 20, 2009. Available from: http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm047470.htm. Accessed Apr 5, 2011.
2. O'Neill RT. A perspective on characterizing benefits and risks derived from clinical trials: can we do more? *Drug Inf J*. 2008;42(3): 235–245.
3. Roth-Cline MD. Clinical trials in the wake of Vioxx: requiring statistically extreme evidence of benefit to ensure the safety of new drugs. *Circulation*. 2006;113(18):2253–2259.

4. Avorn J. Evaluating drug effects in the post-Vioxx world: there must be a better way. *Circulation*. 2006;113(18):2173–2176.

5. Furberg CD, Levin AA, Gross PA, Shapiro RS, Strom BL. The FDA and drug safety: a proposal for sweeping changes. *Arch Intern Med*. 2006;166(18):1938–1942.

6. Guo JJ, Pandey S, Doyle J, Bian B, Lis Y, Raisch DW. A review of quantitative risk-benefit methodologies for assessing drug safety and efficacy – report of the ISPOR risk-benefit management working group. *Value Health*. 2010;13(5):657–666.

7. Garrison LP Jr, Towse A, Bresnahan BW. Assessing a structured, quantitative health outcomes approach to drug risk-benefit analysis. *Health Aff (Millwood)*. 2007;26(3):684–695.

8. Garrison LP. Regulatory benefit-risk assessment and comparative effectiveness research: strangers, bedfellows, or strange bedfellows? *Pharmacoeconomics*. 2010;28(10):855–865.

9. Towse A. Net clinical benefit: the art and science of jointly estimating benefits and risks of medical treatment. *Value Health*. 2010; 13(Suppl 1):S30–S32.

10. Schwartz LM, Woloshin S, Welch HG. Using a drug facts box to communicate drug benefits and harms: two randomized trials. *Ann Intern Med*. 2009;150(8):516–527.

11. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. *Health Econ*. 1996;5(6):513–524.