

# BMJ Open Use of surrogate outcomes in US FDA drug approvals, 2003–2012: a survey

Tsung Yu,<sup>1,2</sup> Yea-Jen Hsu,<sup>3</sup> Kevin M Fain,<sup>1</sup> Cynthia M Boyd,<sup>4</sup> Janet T Holbrook,<sup>1</sup> Milo A Puhon<sup>2</sup>

**To cite:** Yu T, Hsu Y-J, Fain KM, *et al.* Use of surrogate outcomes in US FDA drug approvals, 2003–2012: a survey. *BMJ Open* 2015;**5**:e007960. doi:10.1136/bmjopen-2015-007960

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2015-007960>).

Received 13 February 2015  
Revised 3 October 2015  
Accepted 26 October 2015

## ABSTRACT

**Objective:** To evaluate, across a spectrum of diseases, how often surrogate outcomes are used as a basis for drug approvals by the US Food and Drug Administration (FDA), and whether and how the rationale for using treatment effects on surrogates as predictors of treatment effects on patient-centred outcomes is discussed.

**Study design and setting:** We used the Drugs@FDA website to identify drug approvals produced from 2003 to 2012 by the FDA. We focused on four diseases (chronic obstructive pulmonary disease (COPD), type 1 or 2 diabetes, glaucoma and osteoporosis) for which surrogates are commonly used in trials. We reviewed the drug labels and medical reviews to provide empirical evidence on how surrogate outcomes are handled by the FDA.

**Results:** Of 1043 approvals screened, 58 (6%) were for the four diseases of interest. Most drugs for COPD (7/9, 78%), diabetes (26/26, 100%) and glaucoma (9/9, 100%) were approved based on surrogates while for osteoporosis, most drugs (10/14, 71%) were also approved for patient-centred outcomes (fractures). The rationale for using surrogates was discussed in 11 of the 43 (26%) drug approvals based on surrogates. In these drug approvals, we found drug approvals for diabetes are more likely than the other examined conditions to contain a discussion of trial evidence demonstrating that treatment effects on surrogate outcomes predict treatment effects on patient-centred outcomes.

**Conclusions:** Our results suggest that the FDA did not use a consistent approach to address surrogates in assessing the benefits and harms of drugs for COPD, type 1 or 2 diabetes, glaucoma and osteoporosis. For evaluating new drugs, patient-centred outcomes should be chosen whenever possible. If the use of surrogate outcomes is necessary, then a consistent approach is important to review the evidence for surrogacy and consider surrogate's usage in the treatment and population under study.

## INTRODUCTION

A surrogate outcome is a biomarker or an intermediate outcome that substitutes for patient-centred outcomes, that is, outcomes that patients notice and care about such as

## Strengths and limitations of this study

- This study is one of the first to examine how a national policymaker, in this case, the US Food and Drug Administration (FDA), handles surrogate outcomes when making regulatory decisions.
- For four diseases, we reviewed all drug approvals in 2003–2012. We reviewed, for each drug, the documents of drug labels and medical reviews in order to have a comprehensive assessment of how the treatment effect evidence on surrogate outcomes was considered by drug reviewers.
- We focused on only four chronic diseases and reviewed what was documented by the FDA drug reviews. This limits the generalisability of our findings.

survival, function, symptoms and health-related quality of life.<sup>1,2</sup> Because using patient-centred outcomes in randomised clinical trials (RCTs) may require a study that is larger and takes longer, in certain disease areas, surrogate outcomes are commonly used as the primary outcomes in designing RCTs, to save time, sample size and resources to show a particular treatment effect size.<sup>3</sup> For example, Gandhi *et al*<sup>4</sup> found that, in 436 registered RCTs in type 1 or 2 diabetes, only 78 (18%) trials chose patient-centred outcomes as primary outcomes. Most trials used glycosylated haemoglobin to test the efficacy of diabetes drugs rather than assessing their effects on outcomes that have direct impacts on patients, such as cardiovascular events.

However, there are dangers in relying entirely on surrogate outcomes for treatment effect evidence.<sup>5</sup> Two classic examples are encainide and flecainide, which were new agents approved by the US Food and Drug Administration (FDA) for suppressing ventricular arrhythmias to reduce cardiovascular-related death. Although these agents had an effect on surrogate outcomes (arrhythmias), a clinical trial conducted to evaluate their effect on survival showed that they actually increased the risk of death in patients.<sup>6</sup>



CrossMark

For numbered affiliations see end of article.

## Correspondence to

Dr Tsung Yu;  
tsung.yu@uzh.ch

A metaepidemiological study carried out by Ciani *et al*<sup>7</sup> also found that a larger treatment effect is more often observed in clinical trials using surrogates as primary outcomes than in trials using patient-centred outcomes. Thus, use of surrogate outcomes in RCTs does not provide sufficient clarity for understanding the actual benefits and harms for patients taking the drugs. This poses a real challenge to the regulatory bodies and health technology assessment agencies to make licensing and coverage decisions for prescription drugs.

Although policymakers such as the US FDA commonly face the challenges of relying on surrogate outcomes to make decisions about prescription drug safety and effectiveness, little is known about how such challenges are addressed. The challenges include, first, to properly evaluate the evidence supporting the use of surrogate outcomes ('validity').<sup>1-3</sup> For example, the International Conference on Harmonisation guidelines for the conduct of clinical trials for the registration of drugs (ICH-9) criteria describe a hierarchy of evidence for surrogacy.<sup>8</sup> The evidence for surrogacy may come from pathophysiological studies suggesting the biological plausibility of the association between surrogate outcomes and patient-centred outcomes, or from observational studies demonstrating the association between them. The highest level of evidence requires that RCTs have shown that the treatment effects on surrogate outcomes can predict the treatment effects on patient-centred outcomes. Another challenge for regulatory bodies is when the evidence supporting the use of drugs includes primarily surrogate outcomes (eg, a difference in the biomarker measures between treatment groups), how one can properly make a clinical interpretation of such evidence.

It is not clear if the FDA adopts a consistent approach to the use of surrogate outcomes for drug approvals across a spectrum of diseases. Our study aim was to provide empirical evidence on how surrogate outcomes are handled by the FDA. We reviewed the drug approvals produced by the FDA for four diseases, from 2003 to 2012, to learn how often these approvals were based on surrogate outcomes, and whether and how the rationale for using surrogate outcomes was discussed.

## METHODS

### Selection of drug approvals

We used the Drugs@FDA website (<http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm>) to identify all drug approvals produced from January 2003 to December 2012 (n=1043), by the US FDA. Drugs@FDA is an open access database for drug products approved by the FDA; it contains a drug approval package, including prescribing information, approval letters and FDA reviews such as medical, chemistry, pharmacology and statistical reviews. These reviews provide scientific analysis of a drug product and explain the FDA's thinking for the approval decision.

Two authors (TY and YJH), working independently, screened the list to select the approvals that were eligible.

The inclusion criteria were drug approvals where the drugs are indicated for the treatment of chronic obstructive pulmonary disease (COPD), diabetes, glaucoma or osteoporosis. We focused on these four diseases because surrogate outcomes (lung function parameters for COPD, blood sugar level for diabetes, intraocular pressure (IOP) for glaucoma and bone mineral density for osteoporosis, respectively) are commonly used as primary outcomes in RCTs and all of them are 'well established' surrogates according to the guidance documents issued by the FDA.<sup>9-12</sup> We excluded the drugs that are only indicated for a specific symptom related to the diseases or indicated for a specific patient subpopulation. Thus, we excluded a glaucoma drug that is indicated as an adjunct to ab externo glaucoma surgery, a diabetes drug approved for treating adult patients with endogenous Cushing's syndrome who have type 2 diabetes, and a drug treating diabetic peripheral neuropathic pain. We also removed any duplicate records. If there was a disagreement between the two authors about including or excluding a drug approval, we resolved it by discussion.

### Data extraction

During the drug's approval process, the FDA review team critically evaluates different aspects of the drug's benefits and harms, and produces review documents, including the medical, chemistry, pharmacology and statistical reviews, etc. For each included drug approval, we retrieved the prescribing information and medical reviews that were available on the Drug@FDA website. We focused on medical reviews instead of other reviews because the FDA medical reviews are, as we learned during pilot-testing of data extraction, most likely the review documents where the FDA reviewers address the issue of outcome selection. In addition, the medical review documents provide the FDA reviewers' assessment of clinical evidence that establishes the efficacy and safety of the drug.

We developed and pilot-tested a standardised form for data extraction. Using the documents of prescribing information and medical reviews, we extracted the information on indications and the primary outcomes that the indications were based on. If it was not clear what outcomes the indications were based on, we reviewed the outcomes reported in the clinical studies cited in the prescribing information to make a judgment. We then categorised these outcomes into a surrogate outcome or a patient-centred outcome using the definition mentioned previously. For each drug approved based only on surrogates, we examined if the rationale for using surrogate outcomes was discussed or not (yes/no). We also assessed whether the surrogate was identified as being based on the highest level of evidence for surrogacy using the ICH-9 criteria. Finally, we examined if the reviewers interpreted surrogate outcome results in

RCTs, using metrics such as minimal important difference (MID)<sup>13</sup> or a threshold that has been shown to be linked to patient-centred outcomes. Two authors (TY and YJH) independently reviewed all documents and extracted the data. The discrepancies between authors were resolved through discussion. We used descriptive statistics to summarise our findings.

## RESULTS

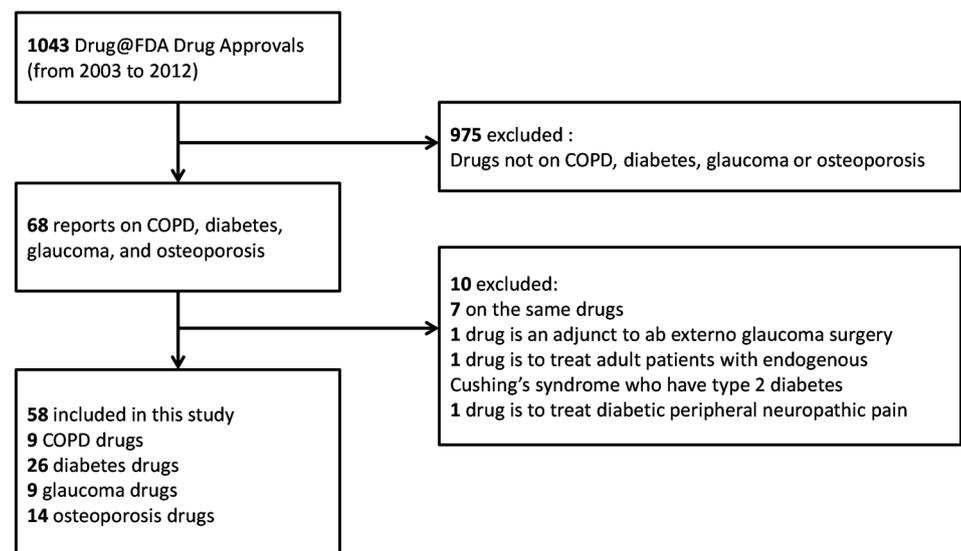
Sixty-eight of 1043 (7%) drug approvals were about COPD, diabetes, glaucoma or osteoporosis, and 58 (58/1043; 6%) of these were eligible for our study. The reasons for exclusion of approvals are summarised in figure 1. Of the 58 included approvals, 9 were for COPD (16%), 26 (45%) for diabetes, 9 (16%) for glaucoma and 14 (24%) for osteoporosis. For three of the four examined conditions, the drug approvals were mostly based only on a surrogate outcome (COPD (7/9 approvals were based only on a surrogate, 78%), diabetes (26/26 approvals, 100%) and glaucoma (9/9 approvals, 100%), see online supplementary table S1). COPD drug approvals were primarily based on the effects on improving lung function, with the exception of two drug approvals (*SPIRIVA HANDIHALER* and *DALIRESP*), which also examined COPD exacerbations. All diabetes drug approvals reviewed were based on lowering blood sugar level and all glaucoma drug approvals reviewed were based on lowering IOP. Most drug approvals for osteoporosis (10/14; 71%) were based on both, surrogate outcomes (bone mineral density) and patient-centred outcome (fractures).

Among the drugs that were approved based only on surrogates, 11 (11/44, 25%) discussed, in the medical review, the rationale for using surrogate outcomes to demonstrate drug efficacy for regulatory approval (table 1). For COPD drug approvals based on surrogates, a medical review for one drug (*TUDORZA PRESSAIR*) mentioned the limitations of using lung function and the

importance of evaluating patient-centred outcomes such as COPD exacerbations. For glaucoma, the reviews for three drugs (*ALPHAGAN P*, *QOLIANA* and *LUMIGAN*) discussed the rationale for using change in IOP for drug approval. These reviews mentioned the association between high IOP and visual function loss but, did not cite evidence from RCTs that an effect on IOP predicts an effect on visual function. For diabetes, we found that the reviews for seven drugs (*APIDRA*, *SYMLIN*, *EXUBERA*, *JANUVIA*, *JANUMET*, *VICTOZA* and *BYDUREON*) discussed the rationale for use of surrogates and three of them (*SYMLIN*, *VICTOZA* and *BYDUREON*) justified choosing glycaemic control as an outcome by clearly stating the evidence that corresponds to the highest level of evidence for surrogacy using the ICH-9 criteria. For example, in the review of *VICTOZA*, the reviewer stated that “HbA1c has excellent reliability, predicts several diabetes-specific complications, and provides the current basis for treatment decisions. Lowering HbA1c reduces microvascular complications in patients with type 1 and type 2 diabetes and possibly macrovascular complications in patients with type 1 diabetes.” They cited evidence from two long-term RCTs in patients with diabetes to justify the use of surrogates.<sup>14 15</sup> We did not observe a change over time from 2003 to 2012, as to what type of outcomes the drug approvals were based on or how they justified the use of surrogates.

Regarding the interpretation of surrogate outcome results in RCTs, 13 reviews (13/44, 30%) discussed the use of MID or threshold. We found that a review for one drug in COPD (*ARCAPTA NEOHALER*) mentioned a MID and reviews for 12 drugs in diabetes (*AVANDARYL*, *SYMLIN*, *DUETACT*, *EXUBERA*, *JANUVIA*, *BYETTA*, *CYCLOSET*, *ONGLYZA*, *KOMBIGLYZE XR*, *VICTOZA*, *TRADJENTA* and *BYDUREON*) mentioned a threshold (number of patients achieving the target haemoglobin A1C level) that is linked to patient-centred outcomes. We did not find the discussion of MID or threshold in the reviews for glaucoma and osteoporosis.

**Figure 1** Review process. COPD, Chronic obstructive pulmonary disease; FDA, Food and Drug Administration.



**Table 1** Rationale for using surrogate outcomes discussed in drug medical reviews (n=11)

Disease	Drug name	Year approved	Rationale for using surrogate outcomes	The rationale is based on: 'treatment effects on the surrogate outcome predict treatment effects on the patient-centred outcome' (highest level of evidence using the International Conference on Harmonisation guidelines for the conduct of clinical trials for the registration of drugs (ICH-9) criteria for surrogacy)	Evidence cited to support the use of surrogate outcome
Chronic obstructive pulmonary disease (COPD)	TUDORZA PRESSAIR	2012	<i>"Overall, the committee's view was that the Applicant's data for the primary end point of trough forced expiratory volume in one second (FEV1) demonstrated statistical significance, and that these results were clinically meaningful... Comments were made that the results for other measures of efficacy (eg, the St George's Respiratory Questionnaire (SGRQ) and COPD exacerbations), while generally not statistically significant, were nonetheless trending in a direction to support the results for the primary end point... Several comments were made regarding the limitations of FEV1-based end points and the importance of evaluating patient-centred outcomes."</i>	No	None
Diabetes (type 1 or 2)	APIDRA	2004	<i>"GHb (note: glycosylated haemoglobin) results were reported as glycated haemoglobin A1c (HbA1c) equivalents and are directly traceable to the Diabetes Control and Complications Trial (DCCT) reference, for which the relationship between mean BG (blood glucose) (measured by HbA1c) and the risk for vascular complications has been established."</i>	Unclear	Diabetes control and complications trial <sup>29</sup>
Diabetes (type 1 or 2)	SYMLIN	2005	<i>"FDA (note: Food and Drug Administration) accepts reduction in HbA1c as a measure of efficacy in trials of new antidiabetic agents. This use of HbA1c as a surrogate end point reflects the finding that long-term reduction of HbA1c decreases the risk of diabetic complications, particularly retinopathy."</i>	Yes	None

Continued

Table 1 Continued

Disease	Drug name	Year approved	Rationale for using surrogate outcomes	The rationale is based on: 'treatment effects on the surrogate outcome predict treatment effects on the patient-centred outcome' (highest level of evidence using the International Conference on Harmonisation guidelines for the conduct of clinical trials for the registration of drugs (ICH-9) criteria for surrogacy)	Evidence cited to support the use of surrogate outcome
Diabetes (type 1 or 2)	EXUBERA	2006	<i>"An ideal trial would use diabetic complications as end points, but the trial size and duration needed for use of such end points would be very large. There is some controversy about whether HbA1c is truly a good marker of the risk for complications of diabetes. However, the correlation of HbA1c with risk for the development of microvascular disease in type 1 diabetics is well-established (Jeffcoate 2004), and thus HbA1c is a good surrogate end point for the trials of inhaled insulin in type 1 diabetics."</i>	No	Jeffcoate 2004 <sup>30</sup>
Diabetes (type 2)	JANUVIA	2006	<i>"HbA1c is generally considered the most reliable surrogate of glycaemic control, and ultimately predicts late chronic complications of T2DM (type 2 diabetes mellitus) microvascular and macrovascular, as demonstrated in the Diabetes Control and Complications Trial (DCCT) and the United Kingdom Prospective Diabetes Study (UKPDS)."</i>	Unclear	Diabetes control and complications trial and UK Prospective diabetes study <sup>14 29</sup>
Diabetes (type 2)	JANUMET	2007	<i>"HbA1c is generally considered the most reliable surrogate of glycaemic control, and ultimately predicts late chronic complications of T2DM (type 2 diabetes mellitus) microvascular and macrovascular, as demonstrated in the Diabetes Control and Complications Trial (DCCT) and the United Kingdom Prospective Diabetes Study (UKPDS)."</i>	Unclear	Diabetes control and complications trial and UK Prospective diabetes study <sup>14 29</sup>
Diabetes (type 2)	VICTOZA	2010	<i>"HbA1c has excellent reliability, predicts several diabetes-specific complications and provides the current basis for treatment decisions (American Diabetes Association</i>	Yes	Diabetes control and complications trial, UK Prospective diabetes study and diabetes control and complications trial/epidemiology of diabetes interventions

Continued

Table 1 Continued

Disease	Drug name	Year approved	Rationale for using surrogate outcomes	The rationale is based on: 'treatment effects on the surrogate outcome predict treatment effects on the patient-centred outcome' (highest level of evidence using the International Conference on Harmonisation guidelines for the conduct of clinical trials for the registration of drugs (ICH-9) criteria for surrogacy)	Evidence cited to support the use of surrogate outcome
Diabetes (type 2)	BYDUREON	2012	2008)... Lowering HbA1c reduces microvascular complications in patients with type 1 and type 2 diabetes (Diabetes Control and Complications Trial Research Group 1993, UK Prospective Diabetes Study (UKPDS) Group 1998) and possibly macrovascular complications in patients with type 1 diabetes (Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Study Research Group 2005)." "HbA1c has excellent reliability, predicts several diabetes-specific complications and provides the current basis for treatment decisions (American Diabetes Association 2006)...Lowering HbA1c reduces microvascular complications in patients with type 1 and type 2 diabetes (Diabetes Control and Complications Trial Research Group 1993, UK Prospective Diabetes Study (UKPDS) Group 1998). There is weaker evidence showing that lowering HbA1c reduces macrovascular complications in patients with type 1 diabetes (Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Study Research Group 2005)."	Yes	and complications (DCCT/EDIC) study <sup>14 29</sup>  Diabetes control and complications trial, UK Prospective diabetes study and diabetes control and complications trial/epidemiology of diabetes interventions and complications (DCCT/EDIC) study <sup>14 29</sup>
Glaucoma	ALPHAGAN P	2005	"Elevated IOP (intraocular pressure) presents a major risk factor in glaucomatous field loss. The higher the level of IOP, the greater the likelihood of optic nerve damage and visual field loss."	No	None
Glaucoma	QOLIANA	2006	"Elevated intraocular pressure is an aetiological factor in glaucomatous cupping. Higher intraocular pressure corresponds	Unclear	None

Continued

Table 1 Continued

Disease	Drug name	Year approved	Rationale for using surrogate outcomes	The rationale is based on: 'treatment effects on the surrogate outcome predict treatment effects on the patient-centred outcome' (highest level of evidence using the International Conference on Harmonisation guidelines for the conduct of clinical trials for the registration of drugs (ICH-9) criteria for surrogacy)	Evidence cited to support the use of surrogate outcome
Glaucoma	LUMIGAN	2010	<p>with a greater frequency of optic nerve damage. Medical therapy for open-angle glaucoma is aimed at lowering the intraocular pressure below a level that is likely to produce further optic nerve damage."</p> <p>"Intraocular pressure (IOP) is currently the accepted standard for establishing the efficacy of ocular hypotensive medications. IOP is a surrogate end point for potential visual function loss."</p>	Unclear	None

## DISCUSSION

Our study findings suggest that the FDA did not use a consistent approach to address surrogate outcomes when reviewing the drug approvals included in this study. In diseases such as COPD, diabetes and glaucoma, we found that RCT evidence relying on surrogate outcomes forms the basis for FDA drug approvals. But for osteoporosis, treatment effects on the surrogate outcome (bone mineral density) and the patient-centred outcome (fractures) were often examined together when regulatory decisions were made. In addition, the rationale for using surrogate outcomes for drug approval was not always discussed. If it was discussed, drug approvals for diabetes were more likely than drug approvals for the other examined conditions to contain a discussion of RCT evidence demonstrating that treatment effects on surrogate outcomes (blood sugar level) predict treatment effects on patient-centred outcomes (macrovascular or microvascular events).

This study also demonstrates that the FDA regulatory pathway for certain diseases still relies heavily on surrogate outcomes. Similarly, a recent survey of prescription drugs conducted by Downing *et al*<sup>16</sup> found that surrogate outcomes were used as the primary outcomes in about 50% of pivotal trials for FDA regulatory approval. The actual treatment effect of many drugs on patients is thus left to be extrapolated from treatment effect on surrogates by clinicians themselves who prescribe the drugs and by patients themselves who take the drugs. For example, most drugs for diabetes are only indicated for 'glycaemic control' rather than indicated for lowering the risk of patient-centred outcomes such as stroke or amputation. To make an extrapolation of the treatment effect clinically, a MID or threshold is often defined. In this case, the target level for glycaemic control is set as haemoglobin A1C <7% in adult patients with type 2 diabetes mellitus,<sup>17</sup> which is a level of haemoglobin A1C that has been linked to a lower risk of microvascular or macrovascular events. However, we should be cautious when using a threshold of this kind. The target level of surrogates may not hold constant across different drugs (drug classes) or different patient groups,<sup>18 19</sup> since surrogates may have a continuous (instead of dichotomous) and other non-linear associations with the corresponding patient-centred outcomes.<sup>20</sup> Ideally, we should have treatment evidence on outcomes that are directly relevant to patients. RCTs should provide us direct evidence on how much of an impact the drugs have on patient-centred outcomes. Decision-makers can then be better informed of the benefits and harms that the drugs cause to patients.

For making market authorisation or coverage decisions, we suggest that policymakers should consider primarily the evidence on patient-centred outcomes. Some may argue that for some diseases it is not always feasible to design and implement RCTs assessing patient-centred outcomes. In fact, this could be the argument that would be made for three diseases examined in this study

(COPD, diabetes and glaucoma), for which the FDA allows relying on surrogate outcomes to approve the drugs. In such situations, when evidence on patient-centred outcomes for a drug is lacking, drug reviews should properly consider the validity of using surrogate outcomes in the specific drug and population of interest. In our survey, we found the rationale for using surrogate outcomes for drug approval was not often discussed in the FDA medical reviews. Even if the rationale was given, certain reviews were not clear about the role of surrogate outcomes and considered them appropriate for RCTs based solely on the assertion that they are the risk factors for patient-centred outcomes. Some reviews for diabetes drugs considered evidence from RCTs demonstrating that the effect on surrogates can predict the effect on patient-centred outcomes but such evidence was from a limited number of trials and was not examined in a systematic way.

We reviewed published guidance<sup>21–23</sup> on surrogate outcomes and make the following suggestions for drug reviewers (or any decision-makers who need to weigh the benefits and harms of treatments) to properly handle surrogate outcomes:

1. Evidence for surrogacy should be based on RCTs evaluating whether the treatment effect on surrogates predicts treatment effect on patient-centred outcomes

Surrogate outcomes are used in RCTs because they can be an indicator or intermediate variable in the disease process and can substitute for patient-centred outcomes. There is often good evidence from epidemiological studies that demonstrate the association between both outcomes. However, to formally validate a surrogate outcome, it is necessary to have evidence from RCTs assessing whether the treatment effect on surrogates consistently predicts the treatment effect on patient-centred outcomes. Prentice developed a statistical criterion for evaluating surrogate outcomes in trials,<sup>24</sup> which requires that surrogate outcomes fully capture the treatment effect on patient-centred outcomes. However, this criterion is seldom met in practice.

Another statistical approach to validate surrogate outcomes is using data from multiple RCTs that assess surrogate outcomes and patient-centred outcomes.<sup>23</sup> One can build a multilevel model to fit data from multiple trials and calculate a trial-level and an individual-level association between treatment effects on both outcomes, or one can calculate the ‘surrogate threshold effect’ to evaluate the evidence for surrogacy.<sup>23</sup> A detailed discussion of statistical methods for validating surrogate outcomes is out of scope for this article but some references are provided here.<sup>23–25</sup>

2. The evidence for surrogacy may be context-specific

The validity of surrogate outcomes can potentially vary by disease, drug (or drug classes) and subpopulation because surrogate outcomes may not mediate the disease pathway in the same way across different

contexts.<sup>18 19</sup> Additionally, drugs can cause benefits or harms to patients through the effect that is independent of surrogate outcomes.<sup>5</sup> Thus, when evaluating existing evidence for surrogacy, we suggest conducting systematic reviews of RCTs and trying to investigate the heterogeneity of the evidence for surrogacy, and consideration of all important outcomes. For reviewing a new drug, it is probably not common that the validity of the surrogates has been already established in the specific treatment or population under review, so an extrapolation of the treatment effect is inevitable. Nonetheless, it is important for drug reviewers to recognise the limitations of making such extrapolations.

3. The role of postmarketing studies should be emphasised if surrogate outcomes are used as a basis for drug approvals

One way to alleviate the threats of relying on surrogate outcomes for drug approval is by requiring long-term postmarketing studies.<sup>26</sup> Rosiglitazone, for example, was approved by the FDA for effectively controlling the blood sugar level in patients with diabetes. However, later meta-analyses have suggested that rosiglitazone is associated with an unexpected higher risk for cardiovascular events.<sup>27</sup> Accordingly, the FDA now requires drug companies manufacturing diabetes drugs to provide data on cardiovascular outcomes and to continue monitoring the drug safety in postmarketing studies, in certain circumstances.<sup>28</sup> As long as the drugs are approved based on surrogate outcomes without knowing their effect on patient-centred outcomes, we will never be certain of their actual beneficial or harmful effects on patients. We emphasise the importance of conducting long-term safety studies.<sup>16 28</sup>

### Limitation of our survey

Our study only focused on the surrogate outcomes used for drug efficacy and did not address surrogate outcomes that substitute for harms. Harmful events are often rare and may take a longer time to develop so that regulatory agencies may be even more dependent on surrogate outcomes for harms regardless of their validity and will require more data beyond RCTs, such as large and long-term postmarketing studies, to assess the harms. We reviewed four diseases where surrogate outcomes are commonly used but did not review diseases such as cancers or HIV, where the use of surrogate outcomes is also prevalent. There may be considerations with regard to the lack of treatment alternatives, so the use of surrogate outcomes is necessary for cancers or HIV drugs to accelerate the regulatory approval process.<sup>1</sup> We did not evaluate the new drug applications that were declined by the FDA because these documents are not publicly available. There may be more explicit analysis of surrogate outcomes in those documents. We focused on medical reviews of the FDA drug approval process since we found that this is where a discussion of surrogate outcomes would most likely be documented, but there is the possibility that it was mentioned

elsewhere in the FDA reviews. Finally, not documenting the rationale for use of surrogate outcomes does not mean that the FDA reviewers did not take it into account when making decisions. However, a documented discussion of the evidence will certainly increase the transparency of the process in which regulatory bodies consider surrogate outcomes for drug approvals.

## CONCLUSIONS

Our survey findings suggest that, for three of the diseases examined (COPD, diabetes and glaucoma), drugs are approved based on their treatment effects on surrogate outcomes, but that the FDA does not use a consistent approach for surrogates in order to evaluate these drug applications. This makes it difficult to assess and interpret their actual clinical effects on outcomes important to patients. For evaluating new drugs, patient-centred outcomes should be used whenever possible. If the use of surrogate outcomes is needed, assessing the validity of surrogate outcomes and considering the surrogate's usage in the treatment and population under study is necessary to inform a drug evaluation.

### Author affiliations

<sup>1</sup>Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, USA

<sup>2</sup>Epidemiology, Biostatistics, and Prevention Institute, University of Zurich, Zurich, Switzerland

<sup>3</sup>Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

<sup>4</sup>School of Medicine, Johns Hopkins University, Baltimore, Maryland, USA

**Contributors** TY and MAP contributed to the study protocol. TY and Y-JH extracted the data. TY and MAP drafted the first version of the report. All the authors revised it critically for important intellectual content, and all the authors read and approved the final manuscript.

**Funding** Support for this project was provided in part by the Doctoral Thesis Research Fund in the Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health.

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

1. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;69:89–95.
2. Methodology Committee of the Patient-Centered Outcomes Research Institute (PCORI). Methodological standards and patient-centeredness in comparative effectiveness research: the PCORI perspective. *JAMA* 2012;307:1636–40.
3. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. *Stat Med* 2012;31:2973–84.
4. Gandhi GY, Murad MH, Fujiyoshi A, *et al.* Patient-important outcomes in registered diabetes trials. *JAMA* 2008;299:2543–9.
5. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;125:605–13.
6. Echt DS, Liebson PR, Mitchell LB, *et al.* Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 1991;324:781–8.
7. Ciani O, Buyse M, Garside R, *et al.* Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemiological study. *BMJ* 2013;346:f457.
8. International Conference on Harmonization. ICH harmonized tripartite guideline. Statistical principles for clinical trials—E9. 1998. [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Efficacy/E9/Step4/E9\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf) (accessed 24 Nov 2014).
9. The Food and Drug Administration. Guidance for industry. Diabetes mellitus: developing drugs and therapeutic biologics for treatment and prevention. 2008. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM071624.pdf> (accessed 24 Nov 2014).
10. The Food and Drug Administration. Guidelines for preclinical and clinical evaluation of agents used in the prevention or treatment of postmenopausal osteoporosis. 1994. <http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/WomensHealthResearch/UCM131206.pdf> (accessed 24 Nov 2014).
11. The Food and Drug Administration. Guidance for industry. Chronic obstructive pulmonary disease: developing drugs for treatment. 2007. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071575.pdf> (accessed 24 Nov 2014).
12. Weinreb RN, Kaufman PL. The glaucoma research community and FDA look to the future: a report from the NEI/FDA CDER glaucoma clinical trial design and endpoints symposium. *Invest Ophthalmol Vis Sci* 2009;50:1497–505.
13. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407–15.
14. Holman RR, Paul SK, Bethel MA, *et al.* 10-year follow-up of intensive glucose control in type 2 diabetes. *N Engl J Med* 2008;359:1577–89.
15. Nathan DM, Cleary PA, Backlund JY, *et al.* Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N Engl J Med* 2005;353:2643–53.
16. Downing NS, Aminawung JA, Shah ND, *et al.* Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005–2012. *JAMA* 2014;311:368–77.
17. American Diabetes Association. Standards of medical care in diabetes—2013. *Diabetes Care* 2013;36(Suppl 1):S11–66.
18. Baker SG, Kramer BS. Surrogate endpoint analysis: an exercise in extrapolation. *J Natl Cancer Inst* 2013;105:316–20.
19. Uhlig K, Leff B, Kent D, *et al.* A Framework for crafting clinical practice guidelines that are relevant to the care and management of people with multimorbidity. *J Gen Intern Med* 2014;29:670–9.
20. Duckworth W, Abraira C, Moritz T, *et al.* Glucose control and vascular complications in veterans with type 2 diabetes. *N Engl J Med* 2009;360:129–39.
21. Bucher HC, Guyatt GH, Cook DJ, *et al.* Users' guides to the medical literature: XIX. applying clinical trial results. A. how to use an article measuring the effect of an intervention on surrogate end points. evidence-based medicine working group. *JAMA* 1999;282:771–8.
22. Institute for Quality and Efficiency in Health Care. Validity of surrogate endpoints in oncology: executive summary. 2011. [https://www.iqwig.de/download/A10-05\\_Executive\\_Summary\\_Surrogate\\_endpoints\\_in\\_oncology.pdf](https://www.iqwig.de/download/A10-05_Executive_Summary_Surrogate_endpoints_in_oncology.pdf) (accessed 24 Nov 2014).
23. Lasserre MN. The Biomarker-Surrogacy Evaluation Schema: a review of the biomarker-surrogate literature and a proposal for a criterion-based, quantitative, multidimensional hierarchical levels of evidence schema for evaluating the status of biomarkers as surrogate endpoints. *Stat Methods Med Res* 2008;17:303–40.
24. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989;8:431–40.
25. Burzykowski T, Buyse M. An alternative measure for meta-analytic surrogate endpoint validation. In: Burzykowski T, Molenberghs G, Buyse M, eds. *The evaluation of surrogate endpoints*. Springer, 2005:323–40.
26. The Food and Drug Administration. Structured approach to benefit-risk assessment in drug regulatory decision-making. Draft PDUFA V implementation plan—February 2013 fiscal years



- 2013–2017. 2013. <http://www.fda.gov/downloads/ForIndustry/UserFees/PrescriptionDrugUserFee/UCM329758.pdf> (accessed 24 Nov 2014).
27. Nissen SE, Wolski K. Rosiglitazone revisited: an updated meta-analysis of risk for myocardial infarction and cardiovascular mortality. *Arch Intern Med* 2010;170:1191–201.
28. The Food and Drug Administration. Guidance for industry: diabetes mellitus—evaluating cardiovascular risk in new antidiabetic therapies to treat type 2 diabetes. 2008. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071627.pdf> (accessed 24 Nov 2014).
29. Nathan DM, DCCT/EDIC Research Group. The diabetes control and complications trial/epidemiology of diabetes interventions and complications study at 30 years: overview. *Diabetes Care* 2014;37:9–16 .
30. Jeffcoate SL. Diabetes control and complications: the role of glycated haemoglobin, 25 years on. *Diabet Med* 2004;21: 657–65.