17-May-2021
BMJ-2021-064952 entitled "Using surrogate endpoints in drug regulation and health technology assessment: Methodological advances and the case for change"

Dear Dr. Naci,

Thank you for sending us this paper and giving us the chance to consider your work. We sent it out for external peer review and discussed it at the Analysis manuscript committee meeting (present: Sophie Cook, Peter Doshi, Emma Rourke, Henry Scowcroft, Paul Simpson).

Unfortunately we do not consider it suitable for publication in its present form. However if you are able to amend it in the light of our and/or reviewers' comments, we would be happy to consider it again.

The reviewers' comments are at the end of this letter.

The editors' comments are listed below:

1. Our overall feeling was that we'd like to offer a chance for a major revision, but we had a lot of trouble with the bit about NICE as it threw us while reading and felt like an abrupt ending. Essentially, it seems like either the NICE guidance becomes the front and center focus of the article, in which case we need to know a lot more about it (what it says and why it matters), or alternatively, it would better to take it out. We suspect that not much can be said about NICE at the moment, so the latter approach of taking it out makes more sense. In that case, you probably need to find a better hook to make the piece feel topical. This was our only major comment - there are a few more minor comments - and a lot of helpful feedback from the reviewers that you will find below.

2. The details on validations methods will likely not be terribly engaging for the majority of readers. Instead, it might be more useful to spend more time explaining an aspect of the argument that is not likely obvious at first blush like why a surrogate endpoint can be considered valid for some drugs but not for other drugs.

3. Would be nice to have a little more scaffolding for readers not fully engaged in this area: what led to surrogates becoming acceptable and more widely used in the approval process? Is it driven by the pharmaceutical companies seeking faster routes to market? Equally if the regulatory bodies said surrogate endpoints weren't acceptable, then the nature of trials would surely have to change? Authors highlight that surrogate endpoints are used in the majority of expedited programs - is there an alternative to using surrogate endpoints in this context that wouldn't increase the time to market?

4. "…sound methods for surrogate validation" - sound methods according to who? What standard are the authors trying to hold the regulators to?

5. is a surrogate endpoint the same as a validated predictor?

6. Your article seems largely accepting of surrogate measures, with the tack that we just need to improve statistical methods. But aren't there clinical scenarios where surrogate endpoints would not be acceptable e.g. acute, short-lived, self-limiting infections? Why would we need surrogates for drugs to treat uncomplicated skin infections? Perhaps you can acknowledge this in the part where you set the scene, so the scope is clearer.

7. When revising, please ensure the article is within 1800-2000 words.

We hope that you will be willing to revise your manuscript and submit it within 4-6 weeks. When submitting your revised manuscript please provide a point by point response to our comments and those of any reviewers.

Please note that resubmitting your manuscript does not guarantee eventual acceptance, and that your revision may be sent again for review.

Please click the the link to start the resubmission process: *** PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. ***

https://mc.manuscriptcentral.com/bmj?URL_MASK=9b90fe0103b4459fb2db4c36ab4f13b2

Alternatively, go to https://mc.manuscriptcentral.com/bmj and login to your Author Center. Click on "Manuscripts with Decisions," and then click on "Create a Resubmission" located next to the manuscript number.

If accepted, your article will be published online at bmj.com, the canonical form of the journal. Please note that only a proportion of accepted analysis articles will also be published in print.

I hope you will find the comments useful. Please don't hesitate to contact me if you wish to discuss this further.

Yours sincerely,
Peter Doshi
Senior Editor, The BMJ
Baltimore, MD, U.S.A.
pdoshi@bmj.com

*** PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. ***

https://mc.manuscriptcentral.com/bmj?URL_MASK=1bbb1f042eb54a9c83f7ee7f1e250c13

IMPORTANT: Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

IMPORTANT INFORMATION TO INCLUDE IN A RESUBMISSION

Instead of returning a signed licence or competing interest form, we require all authors to insert the following statements into the text version of their manuscript:

Licence for Publication
The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ and any other BMJPGL products and sublicences such use and exploit all subsidiary rights, as set out in our licence (http://group.bmj.com/products/journals/instructions-for-authors/licence-forms).

Competing Interest
Please see our policy and the unified Competing Interests form http://resources.bmj.com/bmj/authors/editorial-policies/competing-interests. Please state any competing interests if they exist, or make a no competing interests declaration.

Reviewer(s)' Comments to Author:

Reviewer: 1

Recommendation:

Comments:
Comments for Editor and Author

Comments on Using surrogate endpoints in drug regulation and health technology assessment: Methodological advances and the case for change
Submitted by: Peter Lurie and DeAnna Nara, Center for Science in the Public Interest

Generally, this article is timely and of great importance. The article's title makes the case that the science of surrogate endpoint markers has advanced and seeks to showcase NICE's efforts to address that emerging science. But after some appropriate review and setup, the reader arrives at p. 9 lines 4-10 (for the reviews) and p. 9, lines 27-33 (for the proposals) expecting some details on precisely what NICE has found and is proposing but finds instead some generalities implying that the authors support a more restrictive approach to surrogate markers, but little confirmation that this is so. Clearly, the authors are treading a difficult line between too much and too little information, but the current approach leaves the reader unsatisfied. Perhaps tables providing more detail would help. There is substantial relevant evidence missing from the article in its current draft. There is a large body of papers addressing this subject, but the authors need to do more to convince the reader that their review was systematic.

Pg. Line Comment

4 5/6 The authors' statement is true but doesn't fully capture regulatory agencies' ability to require Phase 4 post marketing studies.
4 15/16 Can the authors point to systematic efforts to assess the frequency of unconfirmed surrogates? See, e.g., https://pubmed.ncbi.nlm.nih.gov/31135808/
4 31 Do you mean that that surrogate markers were the sole basis for approval?
4 51-52 This description does not accurately convey the nature of expedited programs, at least in the US. "Expedited programs" is an umbrella term that includes priority review, fast track, breakthrough therapy and accelerated approval. These are described in detail here: https://www.fda.gov/media/86377/download. Only the last of these inherently involves surrogate markers and then sponsors are supposed to conduct Phase IV studies that document the impact of the drug upon non-surrogate outcomes.
5 7/8 Reference needed for "In 2018, ¾ of new drugs approved by FDA benefitted from at least one expedited program…."
5 37/38 Please provide reference for 2018 FDA table
5 35 Recommend deleting sentence as topic sentence does not align will with rest of paragraph
5 42-7 Can the authors comment on what to them constitutes satisfactory validation?
6 20-26 Are the authors aware of any attempts to determine what percentage of surrogates used to approve drugs actually meet the IQWiG threshold?
6 35/36 Expand on what the variations are
6 51/52 Expand on greater transparency for NICE; what is the transparency for others?
8 10-17 Brought up and included in previous page (lines 42-7)
8 10-27 Authors could consider a table describing all the methodologies for establishing the predictive value of a surrogate marker
8 30/31 "Proposed Changes for surrogate of health technology evaluation". This leads to a summary of the reviews in lines 3-10 on p. 9, but this is at a very high level. Can the authors produce a table that would provide more detail?
9 33/34 Summarize the proposals here in another table would give more granularity to the article; we do not know from this description what these proposals are
10 12-36 Key Messages are too generalized
16 Box 1 Could be deleted; not well conceptualized. Critical points can be incorporated into the text.

Additional Questions:
The BMJ uses compulsory open peer review. Your name and institution will be included with your comments when they are sent to the authors. If the manuscript is accepted, your review, name and institution will be published alongside the article.

If this manuscript is rejected from The BMJ, it may be transferred to another BMJ journal along with your reviewer comments. If the article is selected for publication in another BMJ journal, depending on the editorial policy of the journal your review may also be published. You will be contacted for your permission before this happens.

For more information, please see our peer review terms and conditions.

Please confirm that you understand and consent to the above terms and conditions.: I consent to the publication of this review

Please enter your name: Peter Lurie and DeAnna Nara

Job Title: Executive Director

Institution: Center for Science in the Public Interest

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here:

BMJ are working with ORCID to recognise the importance of the reviewer community. Reviewers are now able to share their activity by connecting their review to their ORCID account to gain recognition for their contributions.

Only the Journal title will be uploaded into the reviewer’s ORCID record, along with the date the record was uploaded; there is no identification of the article’s title or authors. Records are uploaded once a decision (accept, reject, or revision) has been made on the article.

Would you like to be accredited by ORCID for this review?: Yes

Reviewer: 2

Recommendation:

Comments:
This is a well written manuscript on an important topic that is applicable across diseases and conditions. The authors lay out their arguments for the necessity of stricter requirements for the use of surrogate endpoints in regulatory trials of drugs or technologies. The discussion is balanced and clear. What I am struggling with is that it is still unclear what the authors' specific asks are – they mention that "these proposals were included in the consultation on the changes to NICE technology evaluation methods that concluded in December 2020" but the decisions won't be made until September 2021 when the program manual is published. I think the box is a way to clarify this, but as currently written, the box has a lot of text without listing specific criteria. Could the authors make this clearer? How about a box that has current criteria on the left and proposed criteria (specifically delineated) on the right?

Specific Comments:
Event-free survival may not be the best example of a surrogate endpoint. Event-free survival is typically defined as the time the patient remains free of certain complications or events that the treatment was intended to prevent or delay. These events may include the return of the cancer or the onset of certain symptoms, such as bone pain from cancer that has spread to the bone. Although event-free survival may be based on a surrogate, such as radiographic appearance of the cancer, it doesn't have to be. It may be based on patient-important outcomes of pain or symptom control. Better examples of surrogate endpoints could be the laboratory markers in diabetes trials, such as urine microalbumin or glycemic control levels as measured by HbA1c. A table with examples of surrogate and patient important outcomes may be helpful to the readers. Overall survival is important but not always the most important for all patients in all stages of life.

Authors argue that the main benefit of surrogate markers is the speed of the regulatory approval process but – to be fair – it should be mentioned that the use of patient-important outcomes may simply not be feasible in some situations. These situations may include the long time to the actual outcome (e.g., in diabetes, blindness or renal failure) or studies or rare or ultra-rare disorders (in which power cannot be achieved yet progress needs to be made).

One issue with surrogate endpoints that the authors point out is the lack of validation and lack of agreement on the frameworks for validation of surrogate endpoints. Although the authors describe some of these frameworks and methods for assessing validity of surrogate endpoints, there is an inherent problem with this. Even when surrogate endpoints are highly predictive of the outcomes, they may not be predictive of treatment effects. Although authors mention ways to determine the predictive ability of surrogates for treatment effects, applying this across treatments is inherently problematic. Different treatments (including different agents within a drug class) may have unintended consequences or may work through pathways unrelated to the measured surrogates. Although validating surrogates is important, I don't think they can 100% ensure that surrogate endpoints and outcomes will align in any given trial.

Authors mention many European decision-making bodies but do not mention the U.S. FDA or other non-European drug regulatory bodies. This may be included for completeness.

Additional Questions:
The BMJ uses compulsory open peer review. Your name and institution will be included with your comments when they are sent to the authors. If the manuscript is accepted, your review, name and institution will be published alongside the article.

If this manuscript is rejected from The BMJ, it may be transferred to another BMJ journal along with your reviewer comments. If the article is selected for publication in another BMJ journal, depending on the editorial policy of the journal your review may also be published. You will be contacted for your permission before this happens.

For more information, please see our peer review terms and conditions.

Please confirm that you understand and consent to the above terms and conditions.: I consent to the publication of this review

Please enter your name: Kasia J Lipska

Job Title: Associate Professor of Medicine (Endocrinology)

Institution: Yale School of Medicine

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here:

BMJ are working with ORCID to recognise the importance of the reviewer community. Reviewers are now able to share their activity by connecting their review to their ORCID account to gain recognition for their contributions.

Only the Journal title will be uploaded into the reviewer’s ORCID record, along with the date the record was uploaded; there is no identification of the article’s title or authors. Records are uploaded once a decision (accept, reject, or revision) has been made on the article.

Would you like to be accredited by ORCID for this review?: Yes

Reviewer: 3

Recommendation:

Comments:
The title of the paper is somewhat disingenuous. Its focus is on HTA methodology. The authors take a quite narrow view of the role of surrogate endpoints and describe a 'case for change' which they appear to be already actively embarked upon with NICE. There is no up-to-date reporting of what steps regulators such as EMA and FDA are taking with regard to surrogate endpoints. The 'case for change' needs to be made against the background of a wider understanding of clinical research, both as it currently is and the directions in which it is heading.

Although the decision is the Editors' I feel that for a journal like BMJ this paper needs to reflect a wider background and identify the place this work has in that environment. However a paper like this is good for the BMJ's audience. I would like to see a Plain Language Summary added to it.

Having made that criticism I enjoyed reading this paper, and I am not a methodologist. It opened up a rare insight into the world of methodology and statistical analysis. It has a view of a real-world issue but the lack of context means that the paper is unbalanced, lacking a rounded summary of the views of (and work being undertaken by) other stakeholders. The authors do indicate that acceptance of uncertainty (properly contextualised) is appropriate and mention a couple of times that on occasions 'flexibility' is required in decision making. As a patient involved in NICE Technology Appraisals I have seen first-hand such flexibility at work (NICE TAs 179 and 185) and I enjoyed the convoluted language employed to justify it, albeit that was some time ago.

The paper focusses on HTA methodology and change rather than that of the drug regulators, principally FDA and EMA. I would agree that the two main marketing authorisation regulators have weak processes regarding surrogate endpoints. They consult with pharma companies developing a registrational study and accept ahead of the trial which surrogates are used. My understanding is that HTA considerations are not sought at this stage, creating challenges for HTA agencies. This context is not mentioned. This will change in the UK with the ILAP programme managed by MHRA/NICE/SMC. FDA/EMA they do have a systematic approach, the question is how consistently it is applied. (I have been a member of scientific advisory panels for EMA).

The practical problems with surrogate endpoints are legion. The paper, understandably touches on the statistical uncertainties, rather than the clinical uncertainties which underlie them. As patient advocates we like certainty and surrogates do not help. We are aware of more than 20 different descriptions of the PFS endpoint in Patient Consent Information leaflets and are (anecdotally) aware of research clinicians using a different definition from the one in the Consent Information for the study they are working on. The challenge presented by surrogates is well summarised in this paper by the examples of RCT data contradicting data from the Phase II registrational study using a surrogate endpoint. Oloratumab in soft tissue sarcoma is engraved on my advocacy experience.

The issue of regulators requiring confirmation studies is touched upon in the paper but the wider context of 'real-world' studies to provide confirmation of benefit is mentioned almost in passing, nor are the methodological challenges such studies present. The detail is not for this paper but the challenge deserves mention. As the paper points out too often the surrogate is replicated. High quality real-word data is, it must be said, something patient groups are keen to see and more robust methodology would be valued.

Phase II studies with surrogates are now dominating cancer research and are increasing elsewhere. The description in the paper of methods for validating surrogate endpoints are a mathematical presentation which goes (at least) one step beyond my ability to follow but I can appreciate the rigour and the value it offers to the development of robust future procedures. I leave other reviewers to consider the maths. My concerns are about the long-term implementation, and thus the value of this work. The shape of future research will be governed by genetically defined disease rarity, the resulting small size of patient cohorts and an urgency to get drugs approved, this is where patient advocacy has impact. This reality is the key context for the future, it is an unstoppable movement. The solutions in this paper are seen from a statistical/methodological viewpoint, although the importance of 'flexibility' in decision-making is appropriately mentioned, but the implementation awaits NICE guidance.

Patient involvement in clinical trial design and development is a growing field. We know from evaluations in cancer, mental health, arthritis and other diseases that patients have important contributions to add in clinical research. Patients involved in this work usually learn about such issues as endpoints 'on the job'. It is often easy to accept a proposed surrogate without being aware of, let alone considering, the complexities but it is important they get to grips with them and are able to challenge the research teams they work with. There is an opportunity for all regulatory agencies to target these patient groups and support their training.

This brings me to PROs – Patient Reported Outcomes. The paper dismisses HRQoL without discussion. Lets consider HRQoL and PROs together. HRQoL uses questionnaire tools to calculate a score representing a patient's quality-of-life, possibly in a number of domains. PROs are data gathered from patients reporting specific effects of treatment, developing disease, function etc. In isolation these tools offer clinicians and patients information which can support decisions about treatment. There is however lack of comparability, no consistency, frequent missing data, and scarcely concealed bias which means that such data are in the 'flexibility' category for HTA appraisals. The issues are being addressed (SISAQOL, COMET etc). FDA and EMA see a future for these tools in providing co-primary endpoints (alongside a surrogate) in a drug trial, thus offering subjective information to regulators. HTA cannot ignore this, even though I accept that these are still developments they present another context specific to the future of surrogate endpoints for regulatory use which this paper should reflect.

Summary.
A revised title making it clear the work is from an HTA standpoint.
A Plain Language Summary
A significant re-write to:
• Add descriptive information so that the background to the use of surrogate endpoints is more completely contextualised and stakeholder views recorded
• Consider how the 'market' is moving and the role that surrogates will play in future. Discussion of the relevance of the ILAP project may be valuable here.
• Consider how patients could 'toughen up' the challenge to trial designers employing surrogate endpoints at an early stage in development
• Discuss PROs used as co-primary endpoints with surrogates, if only to confirm the methodological challenges which PRO developers must address if this is to happen.

Roger Wilson

Comment to Editors

This paper has a very narrow view. I have no doubt that the rigour they propose for validating surrogate endpoints will add statistical value to regulatory review when surrogates are used but as described in its present form it would limit the paper's reach. We have to recognise that the real-world has contexts which impinge on just about every element of their narrative. Thus I am concerned that the paper will be seen as self-serving, rather than adding value to the continuing discussion about surrogates without revision. As a result I have focused my comments on 'contexts'. My mention of the

ILAP Project, which brings together MHRA, NICE and SMC for a co-ordinated approach to drug regulation, is relevant to the eventual use of the authors' work.

When an established market is disrupted by an incomer the instinct of the traditionally grounded market occupant is to focus efforts and tighten its control. Meanwhile the disruptor innovates, reaches customers in new ways, and develops new products. An outsider might take a similar view on drug regulation although it is an established 'market' with occupants who are impossible to displace. There has been a continuing erosion by a combination of patient demand for access to new drugs, changing study methodology to meet evolving science, shortening timescales, smaller trial cohorts etc and now COVID has been an additional disrupting force. Innovation takes on many guises, rarely building on the familiar, more usually coming from a new direction.

These questions need addressing. We can recognise that rigorous standards should apply to surrogate endpoints but what is the place of that rigour in the value chain from trial design to real world use in a rapidly evolving (disrupted) environment? Is the acknowledged need for 'flexibility' more valuable than statistical rigour? Should methodologists be considering how they can add value to the formal inclusion of 'flexibility' in HTA appraisals without constraining the innovation or originality of such contributions? One such area is PROs. Rather than ignoring them and hoping they will go away (market occupant approach) regulatory bodies and their methodologists should get involved with making them better (innovator approach).

Additional Questions:
The BMJ uses compulsory open peer review. Your name and institution will be included with your comments when they are sent to the authors. If the manuscript is accepted, your review, name and institution will be published alongside the article.

If this manuscript is rejected from The BMJ, it may be transferred to another BMJ journal along with your reviewer comments. If the article is selected for publication in another BMJ journal, depending on the editorial policy of the journal your review may also be published. You will be contacted for your permission before this happens.

For more information, please see our peer review terms and conditions.

Please confirm that you understand and consent to the above terms and conditions.: I consent to the publication of this review

Please enter your name: Roger

Job Title: Wilson

Institution: NCRI Consumer Forum - patient reviewer

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here: None

BMJ are working with ORCID to recognise the importance of the reviewer community. Reviewers are now able to share their activity by connecting their review to their ORCID account to gain recognition for their contributions.

Only the Journal title will be uploaded into the reviewer’s ORCID record, along with the date the record was uploaded; there is no identification of the article’s title or authors. Records are uploaded once a decision (accept, reject, or revision) has been made on the article.

Would you like to be accredited by ORCID for this review?: Yes

Date Sent: 17-May-2021