PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Tracking the impact of research on policy and practice: investigating the feasibility of using citations in clinical guidelines for research evaluation
AUTHORS	David Kryl, Liz Allen, Kevin Dolby, Beverley Sherbon and Ian Viney

VERSION 1 - REVIEW

REVIEWER	Andrew McIvor, Firestone Institute for Respiratory Health, McMaster University
REVIEW RETURNED	20/02/2012

GENERAL COMMENTS	This is a timely review. Important to remember "Guidelines" are not standards of care, they are acknowledged by specialists who provide individualized care but written for the non -specialist or GP. Guidelines Groups try to keep appearance of independence and some of this comes from sheltered organizations who get money from Pharma then 'launder' it to say funded by ..... Society. I agree fully that some people should be acknowledged rather than included as authors however in order for this to succeed Academic Centers should be lobbiedto provide academic credit for this along the same lines as mentorship and should frown on those that still insist that all papers from their institutions have their name on it. In the Respirology world some very prominent examples have come to light recently

REVIEWER	Signed, Ivo Abraham /s/ No competing interests to disclose with regards to this paper.
REVIEW RETURNED	21/02/2012

THE STUDY	This is not a patient-based study hence several of the above items do not apply. Reference 17 is not retrievable but is essential as it described the methodology of the study.
RESULTS & CONCLUSIONS	See my comments below. I recommend publication with minor revisions if the authors limit the presumed scope of the study to a double feasibility analysis. In terms of the original intent of the paper, the design is inherently

	flawed and at best limited to the NHS. If the Editor decides to proceed with publication of the paper as presented (with limited revisions), an editorial comment is most in order.
REPORTING & ETHICS	CONSORT, ethics, and plagiarism do not apply, hence my No answer.
GENERAL COMMENTS	REVIEW OF BMJOPEN-2012-000897 Tracking the impact of research on policy and practice: a protocol for the use of clinical guidelines in research evaluation There is little doubt that this paper merits publication because of its benefit as a double feasibility analysis: (1) the feasibility of bibliometric analysis to judge the scientific utility of a paper to the development of guidelines; and (2) the feasibility of using this information to provide feedback to funders/sponsors about the inclusion (and implicitly, impact) of their sponsored research on guideline development. Hence the recommendation to accept pending revisions that frame the paper within this double feasibility assessment. That being said, if however the authors wish to persist in their fundamental claims about bibliometrics of papers in NICE guidelines being an index of impact on policy and practice, the Editor may wish to solicit an accompanying editorial or comment – basically because the argument is circular within the NICE/NHS context. As long as there is an inherent collusion – in the sense of its Latin etymology of cum-ludere or playing together – of one set of guidelines (NICE) translating into one healthcare system's policy (NHS) and constraining clinical decision-making by prescribers at the point of care, bibliometric analysis will neither provide the methodology nor the evidence for determining the impact on policy and practice. The paper asserts that a link between a report's inclusion in a clinical guideline is an indicator of impact, therefore can be used in impact tracking, and be used a metric by funders/sponsors of research to assess the impact of their research investments. The paper also implies that inclusion in a guideline virtually equates with a contribution to policy and practice. This can be challenged. At the risk over oversimplifying, assume that treatment X receives a positive but that treatment Y received a negative guidance from NICE, and that the NHS concurs with this guidance in terms of coverage of X but not of Y. Within the NICE/NHS context, we can indeed speak of an impact on policy ("X shall" and "Y shall not") and an impact on practice (a patient will be treated with X but not with Y). Both conclusions are secondary and inherent to the NICE/NHS context. First, NHS tends to adopt NICE guidance, essentially because of the (partial) service function of NICE to the NHS. Hence, like many treatments before them across many therapeutic areas, X and Y happened to receive, respectively, a positive and negative evaluation – and ipso facto become policy. This policy decision is not based on the evaluation of all guidelines, only those issued by NICE; virtually equating the guidance (i.e., advice to inform policy) with the policy decision itself. This seeming collusion (Latin etymology: cum-ludere, playing together) of NICE guidance and NHS policy is exclusionary ab initio of other guidelines and could be construed that NICE guidelines are superior to other guidelines (they are very good, but they are so within the cost-effectiveness appraisal framework adopted by NICE – which is not the sole valid

	framework). Second, that the guidance about X and Y (and the research papers that led to them) translates into an impact on practice is undoubtedly true for the NHS population – but not because of prescribers using their best clinical judgment about X and Y at the point of care but because X is covered and Y is not. In sum, there are virtually no degrees of freedom in the chain of "NICE guidance -> NHS coverage policy -> clinical practice" because of the "preferred provider status" of NICE relative to the NHS policy and practice. Just focusing on England and Wales, if one truly wants to examine the impact of a research paper, a research program, and/or a sponsor on policy and practice, NICE and NHS must be uncoupled. At the level of NHS policy development, NICE reports should be seen as the technically and scientifically meritorious guidelines that they are – under consideration of NICE's assumptions, scope, design and methodology, and analysis. Considering NICE's track record, its reports may very well deserve "automatic entry" into the pool of guidelines that the NHS should review to set policy. Yet, the NHS should also assess the technical and scientific merit of other guidelines and hence dilute the weight of NICE guidance. Given the proliferation of guidelines consideration should be limited to guidelines developed by credible organizations who have adopted strict methods of guideline development and provide assurances of scientific independence and financial control. In other words, NICE and other guideline developers should be seen as agents to inform NHS policy analysis, development, and setting; not to virtually be the eventual policy. Perhaps trickier is assuring a measurable impact of a research paper, a research program, and/or a sponsor on practice if we define impact on practice as being the outcome of a clinician's decision at the point of care. If treatment options are constrained by what is covered, clinical practice will be constrained as well – whether the constraint is from NHS coverage in the UK, a health authority in continental Europe or the US, or a private payer in the US. It will be difficult to demonstrate the independent impact of research and sponsorship on clinical practice. In the quest to show bibliometrically the impact on either policy or practice, there are yet two other limiting (if not potentially biasing) factors: (1) whether a study has been published at the time of guideline development; the search strategy, including the engines used and the journals indexed; and (2) the level of evidence deemed minimal and the adoption and weighing of this evidence. Here too dilution by using guidelines other than only the NHS' may be the best approach. It is indeed critical to move the main tenets of the paper beyond the NICE/NHS environment and examine the potential generalizability to other healthcare delivery and healthcare financing systems. Noticeably absent from the paper is the integration of the literature on the developments within the evidence-based medicine (EBM) movement. The authors may wish to review a series of papers published in 2005 in Health Affairs, admittedly with a US bent and with both the benefit and constraint of now being 7 years old. In the lead-off paper, Eddy makes the cogent point that the EBM movement has translated into (lots of) evidence-based guidelines and (increasingly more) evidence-based decision making, but not in evidence-based decision-making about the care of individual

	patients. The critical link is the extent to which guidelines are adopted in daily clinical practice; which Timmermans and Mauck, in a paper in the same issue, describe as mixed and disappointing (not surprisingly so, because the uptake of guidelines is about 50/50, not more than a random process). In a study of anemia management with erythropoiesis stimulating agents in cancer patients with anaemia published in the European Journal of Cancer in 2009, Aapro et al. were (among) the first to establish a direct link between the congruence of an individual patient's care with evidence-based guidelines and subsequently achieving evidence-based outcomes. They showed that evidence-based anaemia management - i.e., practicing in accordance with EORTC guidelines derived from the critical analysis of empirical evidence of a body of research reports, severally and jointly – had a direct, independent impact on haemoglobin outcomes. Why this elaborate argument as part of a manuscript review? If the authors intent is to demonstrate the feasibility of (1) bibliometric analysis to judge the scientific utility of a paper to the development of guidelines, and of (2) using this information to provide feedback to funders/sponsors about the impact of their investments, the authors succeeded with a convincing methodological and technical paper. In contrast, to argue that this approach is generalizable beyond the NICE/NHS environment is not sustained. The approach may indeed work within the NICE/NHS context - but not because of the authors' hypothesis that of a traceable chain from research report to guideline to policy to practice, but because of the collusion of guidelines development, policy, and practice constraints inherent to the NICE/NHS context. Relatedly, it may provide UK funders/sponsors with a method for assessing whether studies in their funding portfolio are accepted in NICE guidelines; and, yes, the collusion will give an illusion of impact. In keeping with BMJ policy and my personal practice of unblinding myself as a reviewer if authors are unblinded to reviewers: Ivo Abraham, Professor, Center for Health Outcomes and Pharmacoeconomic Research, University of Arizona, Tucson, AZ, USA. Email: abraham[at]pharmacy,arizona,edu (replace [at] with @ and commas with dots) **REVIEW OF BMJOPEN-2012-000897** **Tracking the impact of research on policy and practice: a protocol for the use** **of clinical guidelines in research evaluation** There is little doubt that this paper merits publication because of its benefit as a double feasibility analysis: (1) the feasibility of bibliometric analysis to judge the scientific utility of a paper to the development of guidelines; and (2) the feasibility of using this information to provide feedback to funders/sponsors about the inclusion (and implicitly, impact) of their sponsored research on guideline development. Hence the recommendation to accept pending revisions that frame the paper within this double feasibility assessment. That being said, if however the authors wish to persist in their fundamental claims about bibliometrics of papers in NICE guidelines being an index of impact on policy and practice, the Editor may wish to solicit an accompanying editorial or comment – basically because the argument is circular within the NICE/NHS context. As long as

	there is an inherent collusion – in the sense of its Latin etymology of *cum-ludere* or playing together – of one set of guidelines (NICE) translating into one healthcare system's policy (NHS) and constraining clinical decision-making by prescribers at the point of care, bibliometric analysis will neither provide the methodology nor the evidence for determining the impact on policy and practice. The paper asserts that a link between a report's inclusion in a clinical guideline is an indicator of impact, therefore can be used in impact tracking, and be used a metric by funders/sponsors of research to assess the impact of their research investments. The paper also implies that inclusion in a guideline virtually equates with a contribution to policy and practice. This can be challenged. At the risk over oversimplifying, assume that treatment X receives a positive but that treatment Y received a negative guidance from NICE, and that the NHS concurs with this guidance in terms of coverage of X but not of Y. Within the NICE/NHS context, we can indeed speak of an impact on policy ("X shall" and "Y shall not") and an impact on practice (a patient will be treated with X but not with Y). Both conclusions are secondary and inherent to the NICE/NHS context. First, NHS tends to adopt NICE guidance, essentially because of the (partial) service function of NICE to the NHS. Hence, like many treatments before them across many therapeutic areas, X and Y happened to receive, respectively, a positive and negative evaluation – and *ipso facto* become policy. This policy decision is not based on the evaluation of all guidelines, only those issued by NICE; virtually equating the guidance (i.e., advice to *inform* policy) with the policy decision itself. This seeming collusion (Latin etymology: *cum-ludere*, playing together) of NICE guidance and NHS policy is exclusionary *ab initio* of other guidelines and could be construed that NICE guidelines are superior to other guidelines (they are very good, but they are so within the cost-effectiveness appraisal framework adopted by NICE – which is not the sole valid framework). Second, that the guidance about X and Y (and the research papers that led to them) translates into an impact on practice is undoubtedly true for the NHS population – but not because of prescribers using their best clinical judgment about X and Y at the point of care but because X is covered and Y is not. In sum, there are virtually no degrees of freedom in the chain of "NICE guidance -> NHS coverage policy -> clinical practice" because of the "preferred provider status" of NICE relative to the NHS policy and practice. Just focusing on England and Wales, if one truly wants to examine the impact of a research paper, a research program, and/or a sponsor on policy and practice, NICE and NHS must be uncoupled. At the level of NHS policy development, NICE reports should be seen as the technically and scientifically meritorious guidelines that they are – under consideration of NICE's assumptions, scope, design and methodology, and analysis. Considering NICE's track record, its reports may very well deserve "automatic entry" into the pool of guidelines that the NHS should review to set policy. Yet, the NHS should also assess the technical and scientific merit of other guidelines and hence dilute the weight of NICE guidance. Given the proliferation of guidelines consideration should be limited to guidelines developed by credible organizations who have adopted strict methods of guideline development and provide assurances of scientific independence and financial control. In other words, NICE and other guideline developers should be seen as agents to *inform* NHS policy analysis, development, and setting; not to virtually *be* the eventual policy. Perhaps trickier is assuring a

	measurable impact of a research paper, a research program, and/or a sponsor on practice if we define impact on practice as being *the outcome of a clinician's decision at the point of care*. If treatment options are constrained by what is covered, clinical practice will be constrained as well –whether the constraint is from NHS coverage in the UK, a health authority in continental Europe or the US, or a private payer in the US. It will be difficult to demonstrate the *independent* impact of research and sponsorship on clinical practice. In the quest to show bibliometrically the impact on either policy or practice, there are yet two other limiting (if not potentially biasing) factors: (1) whether a study has been published at the time of guideline development; the search strategy, including the engines used and the journals indexed; and (2) the level of evidence deemed minimal and the adoption and weighing of this evidence. Here too dilution by using guidelines other than only the NHS' may be the best approach. It is indeed critical to move the main tenets of the paper beyond the NICE/NHS environment and examine the potential generalizability to other healthcare delivery and healthcare financing systems. Noticeably absent from the paper is the integration of the literature on the developments within the evidence-based medicine (EBM) movement. The authors may wish to review a series of papers published in 2005 in Health Affairs, admittedly with a US bent and with both the benefit and constraint of now being 7 years old. In the lead-off paper, Eddy makes the cogent point that the EBM movement has translated into (lots of) evidencebased guidelines and (increasingly more) evidence-based decision making, but not in evidence-based decision-making about the care of individual patients. The critical link is the extent to which guidelines are adopted in daily clinical practice; which Timmermans and Mauck, in a paper in the same issue, describe as mixed and disappointing (not surprisingly so, because the uptake of guidelines is about 50/50, not more than a random process). In a study of anemia management with erythropoiesis stimulating agents in cancer patients with anaemia published in the European Journal of Cancer in 2009, Aapro et al. were (among) the first to establish a direct link between the congruence of an individual patient's care with evidencebased guidelines and subsequently achieving evidence-based outcomes. They showed that evidence-based anaemia management - i.e., practicing in accordance with EORTC guidelines derived from the critical analysis of empirical evidence of a body of research reports, severally and jointly – had a direct, independent impact on haemoglobin outcomes. Why this elaborate argument as part of a manuscript review? If the authors intent is to demonstrate the feasibility of (1) bibliometric analysis to judge the scientific utility of a paper to the development of guidelines, and of (2) using this information to provide feedback to funders/sponsors about the impact of their investments, the authors succeeded with a convincing methodological and technical paper. In contrast, to argue that this approach is generalizable beyond the NICE/NHS environment is not sustained. The approach may indeed work within the NICE/NHS context - but not because of the authors' hypothesis that of a traceable chain from research report to guideline to policy to practice, but because of the collusion of guidelines development, policy, and practice constraints inherent to the NICE/NHS context. Relatedly, it may provide UK funders/sponsors with a method for assessing whether studies in their funding portfolio are accepted in NICE guidelines; and, yes, the collusion will give an illusion of impact. In keeping with BMJ policy

	and my personal practice of unblinding myself as a reviewer if authors are unblinded to reviewers: Ivo Abraham, rofessor, Center for Health Outcomes and Pharmacoeconomic Research, University of Arizona, Tucson, AZ, USA. Email: abraham[at]pharmacy,arizona,edu (replace [at] with @ and commas with dots)

VERSION 1 – AUTHOR RESPONSE

I have made all requested amendments which have been agreed by all of the authors.

One point that I would like to pick up on from the reviewers comments is the perception that the authors believe there is a direct link between citation in guidelines and impact. We agree with the reviewer that this is a correlation to potential impact, rather than an absolute link. I have tried to clarify this in the manuscript, and that the aim is to look at these potential pathways to impact.

Reference number 17 is not yet available publicly, this is the report that the funders commissioned from RAND which is the foundation of this paper. Our intention is to publish this briefing document on the RAND website as and when this article is published. I have contacted RAND and they are in agreement with this approach, they await notification of the appropriate date. Could you please give me a estimate of when this will be, and then an exact date nearer the time.

We have also added a sentence into the discussion with regard to the potential for open global unique researcher identifier systems (such as ORCID - Open and Researcher Contributor ID, www.orcid.org) which will to support acknowledgement and attribution, and therefore greatly help with the tracking and attribution issues raised in the article. Initiatives such as this have started since we drafted the paper but are very pertinent to the discussion.