PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	How innovative are new drugs launched in the UK? A retrospective study of new drugs listed in the British National Formulary (BNF) 2001-2012
AUTHORS	Ward, Derek; Slade, Angharad; Genus, Tracey; Martino, Orsolina; Stevens, Andrew

VERSION 1 - REVIEW

REVIEWER	Joel Lexchin York University, Canada
REVIEW RETURNED	12-Aug-2014

GENERAL COMMENTS	This manuscript explores the level of innovation represented by new chemical entities introduced into the British market between 2001 and 2012. The level of innovation is based on criteria developed by Aronson and the use of these criteria is the major problem that I have with this paper. To begin with, the authors explicitly state that they are only looking at NCEs but the Aronson criteria deal not only with NCEs but also with new indications for older drugs, improved identification of those likely to benefit or be harmed from drugs and modified-release formulations. The authors don't explain which of the criteria they are applying to determine how innovative NCEs are. Furthermore, the review by Kesselheim et al that they cite states "we believe that therapeutic value measures hold the greatest promise for evaluating the effectiveness of investments in drug development by public and private sources." How are the Aronson criteria used to specifically assess therapeutic value? Finally, what sources of information were used to gather the evidence in order to apply the criteria? As it stands it seems that innovation was assessed based on the authors' "gestalt". I also have some further more specific concerns: 1. It's not clear if the authors are excluding things such as extended release formulations and new combinations of existing products. 2. Page 8, line 12: State what level of agreement the Kappa score represents. 3. Page 9, lines 34-38: The authors should refer to the letter from Morgan et al (BMJ 2012;345:5880) that shows that the fall off in NCE introduction was a result of a decrease in late follow-on drugs. 4. Page 10, line 35: What kind of innovation are the initiatives designed to reward?

REVIEWER	Aaron Kesselheim

	Brigham and Women's Hospital/Harvard Medical School
REVIEW RETURNED	13-Aug-2014

GENERAL COMMENTS	In this study, Ward et al analyze the innovativeness of new drugs launched in the UK from 2001-2012 using a previously published framework. They found that the overall proportion of highly innovative drugs was only 26% for the entire 12 year period, and that there was no discernable time trend for this category during the study period. In addition, they found a slight uptrend in the number of slightly innovative drugs since 2004, which mirrored the trend for number of approved drugs overall. As a general matter, I would like more discussion of the difference between a drug being introduced into the UK market and being covered by NHS. Are more innovative drugs covered by the NHS? Is that serving as a reasonable filter? Results Page 8, line 56: I interpreted the sentence as meaning that the "malignant disease and immunosuppression" and "skin" therapeutic areas were kept as separate categories when comparing the proportion of highly innovative products? Is this correct? The p-value should be included in the text. Also, I would suggest comparing the proportion of highly innovative drugs in these therapeutic areas to the proportion coded as highly innovative in all other categories, rather than the total proportion. Discussion: Page 9, line 40: The authors mention here that 5 broad categories accounted for the majority of new drug introductions, but in the Results section, they mention that 6 areas represented 3/4 of all new drugs. Please clarify. Page 10, line 7: The authors mention here a significant limitation of almost all classification systems for pharmaceutical innovativeness, mainly that they don't take account of the public health and health service impact of new drugs. Are there any steps that should be taken to address this? Page 10, line 33: The authors present in this paragraph a comprehensive layout of initiatives taken at the UK and European level to encourage the development of highly innovative drugs. They also point out the long lead-times to impact for these policies. Can the authors propose additional solutions to tackling this drought of innovation? Perhaps from policies currently being implemented by foreign regulatory agencies, including the US Food and Drug Administration? In the Discussion, some mention of the study limitations should be included. For example, given the variety of innovation rating schemes out there, I wonder how much the authors' results would change based on using other rating mechanism? Supplemental data file: Please indicate how the numerical ratings correspond with the degree of innovativeness in Table 1.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

1. This manuscript explores the level of innovation represented by new chemical entities introduced into the British market between 2001 and 2012. The level of innovation is based on criteria developed by Aronson and the use of these criteria is the major problem that I have with this paper. To begin with, the authors explicitly state that they are only looking at NCEs but the Aronson criteria deal not only with NCEs but also with new indications for older drugs, improved identification of those likely to benefit or be harmed from drugs and modified-release formulations.

The criteria developed by Aronson et al (2012) can be used to assess innovation in any medicinal product, including new indications for existing licensed drugs. For example Aronson et al offer aspirin for the prevention of stroke as a successful example of a highly innovative and novel application for an existing compound. However, we investigated only new drugs (new chemical entities [NCEs] and new biological agents) as we were interested in the introduction of novel active substances into the UK medicines market. Our definition of new drugs (and hence those eligible for inclusion in this study) are given in the supplementary file. Modified-release formulations or new combination products that incorporate only existing licensed compounds or biological products were not considered to be new drugs according to our definition. We fully acknowledge that these restrictions mean the focus of the article is narrow. However, we feel that it makes an important contribution to the ongoing debate on the levels of pharmaceutical innovation and offers a particular UK perspective, which is lacking in the current evidence base.

2. The authors don't explain which of the criteria they are applying to determine how innovative NCEs are.

Aronson et al (2012) propose a classification of innovation based on clinical usefulness (a hierarchy of i. benefit in a condition with no existing effective treatment, ii. improvement in the treatment of a condition that does not have a consistently satisfactory treatment, iii. safer treatment, iv. more cost-effective treatment, and v. more convenient treatment) and the nature or degree of innovativeness (process through which the innovation arises, which incorporates pharmaceutical novelty). We used all criteria except clinical usefulness that relates to cost-effectiveness (which was not assessed) and innovativeness that relates to novel applications for existing compounds (which were excluded from our study).

Each new drug was first assessed against the criteria for clinical benefit and then against criteria for the nature of the innovation (see our response to point 4 below) to produce a final categorisation. The criteria assessing clinical benefit partly map to those assessing the nature of innovation, so that an effective drug for a condition with no current treatment will inevitably be classified as highly innovative, improvement in the treatment of condition with no satisfactory existing treatment will be classified as moderately or highly innovative depending on the nature of the innovation, but a more convenient treatment only will not be classified as highly innovative.

We agree that our description of the criteria used in our study to characterise levels of innovation is incomplete. We have amended the Abstract (para 3 lines 4-5), Article Summary (para 2), Introduction (para 1 lines 18-26 and para 3 lines 10-11) and Methods (para 2 lines 3-13) sections to more fully reflect this point, as well as amending Table 1, which describes the criteria that we used in detail. Also see also our response to point 4 below.

3. Furthermore, the review by Kesselheim et al that they cite states "we believe that therapeutic value measures hold the greatest promise for evaluating the effectiveness of investments in drug development by public and private sources." How are the Aronson criteria used to specifically assess therapeutic value?

Therapeutic value in the Aronson et al (2012) criteria are encompassed within the assessment of clinical benefit (see point 2 above), though it may be argued that the term 'clinical benefit' goes beyond a definition of 'therapeutic value', so the text has been amended so that the term clinical benefit is used in preference when referring specifically to the Aronson criteria. Our response to point 2 describes how the manuscript has been revised to better explain how the Aronson criteria were used in practice.

4. Finally, what sources of information were used to gather the evidence in order to apply the criteria? As it stands it seems that innovation was assessed based on the authors' "gestalt".

The authors determined clinical benefit and the nature of the innovation through a free text search of the NHS Evidence web portal (which includes all UK national and selected regional health technology assessment and medicines management resources, including NICE and the BNF), the European Medicines Agency website, and commercial pharmaceutical R&D databases. Where a new drug could not be fully categorised according to either criteria relating to clinical benefit and/or the nature of innovation, an additional search of the published literature (Medline) was conducted, though this was required in only a small minority of cases. The methods section has been amended to describe our approach and the information sources used (see also our response to point 2 above).

5. It's not clear if the authors are excluding things such as extended release formulations and new combinations of existing products.

See our response to point 1 above.

6. Page 8, line 12: State what level of agreement the Kappa score represents.

In the initial assessment, the two analysts agreed on the coding of 210 out of 290 new drugs (72.4%). This is stated in Results para 1 line 3 alongside the kappa statistic for inter-rater agreement.

7. Page 9, lines 34-38: The authors should refer to the letter from Morgan et al (BMJ 2012;345:5880) that shows that the fall off in NCE introduction was a result of a decrease in late follow-on drugs.

The letter from Morgan et al offers a useful insight into the reasons why numbers of new drugs being approved by the FDA fell from the mid to late-1990s onwards (data are presented up to 2010). This suggests that the declining trend was driven by numbers of 'follow-on' products, while numbers of 'first-of-kind' products remained fairly constant. While the overall temporal pattern is slightly different from the UK and European perspective (where annual numbers of new drugs being launched started to increase from 2004 onwards), the apparently consistent level of the more innovative categories of new drugs mirrors our findings. Our study mainly considers the subsequent period, during which time annual numbers of new drug launches were recovering, but the insights offered by this US analysis have been incorporated into the Introduction (para 1 lines 4-6).

Reviewer: 2

8. In this study, Ward et al analyze the innovativeness of new drugs launched in the UK from 2001-2012 using a previously published framework. They found that the overall proportion of highly innovative drugs was only 26% for the entire 12 year period, and that there was no discernible time trend for this category during the study period. In addition, they found a slight uptrend in the number of slightly innovative drugs since 2004, which mirrored the trend for number of approved drugs overall. As a general matter, I would like more discussion of the difference between a drug being introduced into the UK market and being covered by NHS. Are more innovative drugs covered by the NHS? Is that serving as a reasonable filter?

As we were interested in new drug launches in the UK (medicines which are licensed and marketed), we chose the BNF as our data source, as this offers one of the very few comprehensive lists of drugs available for use in the UK that covers the entire time period of study (unlike in the USA, information from regulators is not comprehensive or readily available for the UK; European Medicines Agency data does not include drugs licensed via the UK Medicines and Healthcare Products Regulatory Agency, which form a relatively large proportion of drugs licensed early in the time period of study). The availability of drugs on the NHS is a separate issue to that of new drug launches (though the vast majority of drugs in the BNF are available on the NHS), we now realise that the mention of coverage is potentially confusing to the reader. We have therefore amended the text (Methods para 1 line 4) to omit reference to issues of NHS coverage.

9. Page 8, line 56: I interpreted the sentence as meaning that the "malignant disease and immunosuppression" and "skin" therapeutic areas were kept as separate categories when comparing the proportion of highly innovative products? Is this correct? The p-value should be included in the text.

This interpretation is correct, reflecting how broad therapeutic groups are organised in the BNF. P-values have been included in the relevant text in Results (para 3 lines 7, 10 and 12 – also see our response to point 10 below).

10. Also, I would suggest comparing the proportion of highly innovative drugs in these therapeutic areas to the proportion coded as highly innovative in all other categories, rather than the total proportion.

We agree that this is the preferred analytical approach and while it has not led to a change in those therapeutic areas considered to have an increased proportion of highly innovative or only slightly innovative drugs, the specific p-values have changed and have been includes in the relevant text in Results (see our response to point 9 above).

11. Page 9, line 40: The authors mention here that 5 broad categories accounted for the majority of new drug introductions, but in the Results section, they mention that 6 areas represented 3/4 of all new drugs. Please clarify.

The statement in Discussion (para 1 lines 7-8) is incorrect and has been amended to be in accordance with the statement in the Results (para 3 lines 1-3).

12. Page 10, line 7: The authors mention here a significant limitation of almost all classification systems for pharmaceutical innovativeness, mainly that they don't take account of the public health and health service impact of new drugs. Are there any steps that should be taken to address this?

Our comments in the discussion are meant to highlight this potential limitation of all existing schema for classifying innovation in medicinal products, and an assessment of these factors goes well beyond our study data and analysis. However, we have highlighted this as an area for further research in the Discussion (para 4 lines 7-8).

13. Page 10, line 33: The authors present in this paragraph a comprehensive layout of initiatives taken at the UK and European level to encourage the development of highly innovative drugs. They also point out the long lead-times to impact for these policies. Can the authors propose additional solutions to tackling this drought of innovation? Perhaps from policies currently being implemented by foreign regulatory agencies, including the US Food and Drug Administration?

Many of the initiatives listed either broadly mirror those taken in the USA (for example, increasing use

of adaptive licensing approaches and fast-tracking by regulators, incentives for orphan drug development, funding of translational research programmes) or are independent of location (e.g. focus on increasing academic-commercial collaboration, improved trial designs). Others are more applicable in the UK and Europe given the different payer systems and importance placed on decisions by national and regional health technology assessment bodies. We did not specifically include approaches that extend market-exclusivity (patent protection), and these have been added to the Discussion (para 3 lines 2-3), with reference to Kesselheim et al (2010) and the European Medicines Agency website.

Given the balance to be struck between commercial factors and the public good, we believe the initiatives listed are broadly the right ones at this time, and it is clear that their effectiveness will not be apparent for some years to come. However, we do make reference to the need for future evaluation of these initiatives and have added comments (Discussion para 4 lines 5-12) to highlight the need for better clarity in understanding what payers (public sector policymakers in the UK) and patients really value in pharmaceutical innovation, so that the success of these initiatives can be comprehensively and robustly assessed (also see our response to point 12 above).

14. In the Discussion, some mention of the study limitations should be included. For example, given the variety of innovation rating schemes out there, I wonder how much the authors' results would change based on using other rating mechanism?

We have incorporated mention of study limitations into the Discussion (para 2 lines 3-6 and 21), including mention of the potential for misclassification due to the nature of the value judgements required in making the assessments and the late emergence of a drug's benefits or harms, as well as the inclusion of only newly marketed drugs, which may mean we miss other important areas of innovation.

15. Supplemental data file. Please indicate how the numerical ratings correspond with the degree of innovativeness in Table 1.

The numerical ratings correspond to 1 – slightly innovative, 2 – moderately innovative, and 3 – highly innovative. The supplementary data file has been amended to indicate this.

VERSION 2 – REVIEW

REVIEWER	Joel Lexchin, Professor York University Canada
REVIEW RETURNED	11-Sep-2014

GENERAL COMMENTS	The authors have satisfied most of my initial concerns but I still have questions about how they are characterizing innovation. For instance, the thioglitazone drugs act through a novel pathway. I assume that pioglitazone received a rating of 1 (= slightly innovative) because rosiglitazone was already on the market but what rating would rosiglitazone have received. This is an important question because ultimately both of the glitazones proved to be less beneficial than oral hypoglycemic already on the market because of their safety issues. I also would like more explanation from the authors about how they determined the degree of therapeutic improvement in the cases where a new drug was approved on the basis of surrogate outcomes. Finally, the authors appear to be simultaneously using two distinct reference styles – numerical and author/date.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

1. The authors have satisfied most of my initial concerns but I still have questions about how they are characterizing innovation. For instance, the thioglitazone drugs act through a novel pathway. I assume that pioglitazone received a rating of 1 (= slightly innovative) because rosiglitazone was already on the market but what rating would rosiglitazone have received. This is an important question because ultimately both of the glitazones proved to be less beneficial than oral hypoglycemic already on the market because of their safety issues. I also would like more explanation from the authors about how they determined the degree of therapeutic improvement in the cases where a new drug was approved on the basis of surrogate outcomes.

Response:

The example of pioglitazone being rated as only ‘slightly innovative’ because it was neither the first thioglitazone launched in the UK or offered significant additional clinical benefit over existing launched examples of this class is absolutely correct. Our approach was to consider clinical benefit and likely impact on the existing treatment pathway at the time of UK launch. This means that later identified harms (as well as later identified benefits) would not be reflected in our classification. To make these points clear we have amended the methods (paragraph 2, sentence 1) and discussion (paragraph 2, sentence 3) sections to include a note that the assessment relates to the time of launch and that this is a potential limitation of our study.

In addition, our assessment of clinical benefit allocates drugs into very broad categories based on how they impact patients and the existing treatment pathway. This is essentially a qualitative rather than a quantitative exercise, and inevitably involves value judgements incorporating a broad definition of clinical benefit. It does not include a detailed assessment of the evidence available at the time of launch, and hence whether it derives from direct measurement of the desired outcome or relies on a surrogate outcome. In the example given above, using the Aronson criteria, at launch the thioglitazones may have been considered as providing either “Potential improvement in the treatment

of a condition that does not have a consistently satisfactory treatment" or "More convenient treatment" based on their likely place in the treatment pathway (i.e. for patients whose diabetes is not adequately controlled by diet and existing oral hypoglycaemic medications, but before the addition of injectable insulins). We have previously highlighted the qualitative and potentially subjective nature of our approach in the discussion (paragraph 2, sentence 3), but have now also added a comment mentioning the very broad definition of clinical benefit to sentence 1 of the same paragraph.

2. Finally, the authors appear to be simultaneously using two distinct reference styles – numerical and author/date.

Response:
Thank you for highlighting this. We have amended the manuscript text and table 1 for consistency.