

Methods for using Bing's AI-powered search engine for data extraction for a systematic review

James Edward Hill  | Catherine Harris | Andrew Clegg

Synthesis, Economic Evaluation and Decision Science (SEEDS) Group, University of Central Lancashire, Preston, UK

Correspondence

James Edward Hill, Synthesis, Economic Evaluation and Decision Science (SEEDS) Group, University of Central Lancashire, Fylde Rd, Preston PR1 2HE, UK.
Email: jehill1@uclan.ac.uk

Funding information

NIHR ARC NWC

Abstract

Data extraction is a time-consuming and resource-intensive task in the systematic review process. Natural language processing (NLP) artificial intelligence (AI) techniques have the potential to automate data extraction saving time and resources, accelerating the review process, and enhancing the quality and reliability of extracted data. In this paper, we propose a method for using Bing AI and Microsoft Edge as a second reviewer to verify and enhance data items first extracted by a single human reviewer. We describe a worked example of the steps involved in instructing the Bing AI Chat tool to extract study characteristics as data items from a PDF document into a table so that they can be compared with data extracted manually. We show that this technique may provide an additional verification process for data extraction where there are limited resources available or for novice reviewers. However, it should not be seen as a replacement to already established and validated double independent data extraction methods without further evaluation and verification. Use of AI techniques for data extraction in systematic reviews should be transparently and accurately described in reports. Future research should focus on the accuracy, efficiency, completeness, and user experience of using Bing AI for data extraction compared with traditional methods using two or more reviewers independently.

KEYWORDS

AI, artificial intelligence, Bing, data extraction, machine learning, systematic review

Highlights

What is already known

- Data extraction in systematic reviews take a substantial amount of time to undertake.
- Natural language modelling may provide possible methods for enhancing data extraction processes.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Research Synthesis Methods* published by John Wiley & Sons Ltd.

What is new

- Our paper shows how to use Bing AI to facilitate in data extraction.
- Our paper also discusses the strengths and weaknesses of this approach and recommendations for future development of research.

Potential impact for *Research Synthesis Methods* readers

- Our paper provides a standard method which would facilitate in standardising the use of Bing AI for data extraction.
- Using Bing AI for data extraction may provide enhancements regarding increase accuracy and in timesaving.

1 | INTRODUCTION

Systematic reviews are inherently time-consuming and resource-intensive.¹ It reflects the rigorous and comprehensive nature of the processes involved, as well as the requirement for transparency to allow appropriate scrutiny and replication.^{2,3} The process of manually extracting relevant data from numerous research articles is a stage that necessitates a considerable amount of reviewer involvement.¹ Reviewers need to carefully read and analyse each paper, identify relevant information, and extract data according to predetermined criteria.⁴ Despite this large time-burden, errors in this process can be quite common and are often missed.^{5,6} These errors can lead to substantial issues and in some cases lead to notable changes in the effect estimates presented.⁶

Natural language processing (NLP) artificial intelligence (AI) techniques hold significant potential in revolutionising data extraction for systematic reviews.⁷ NLP enables computers to understand and interpret human language, making it an ideal tool for extracting relevant data from a vast corpus of textual information.⁸ By leveraging NLP algorithms, researchers can automate the extraction process, significantly reducing the manual effort required.¹ NLP algorithms can identify and extract key information such as study characteristics, outcomes, and statistical data from research papers.^{1,9} By harnessing the power of NLP, systematic reviewers may save substantial time and resources, accelerate the review process, and enhance the overall quality and reliability of the extracted data.^{9,10} One of these models which has been given a large amount of public attention and may provide opportunities to enhance the data extraction process is ChatGPT.^{11–13} However, despite the potential opportunities there have been many concerns regarding how and when ChatGPT should be used to facilitate systematic reviews.¹² This field is constantly evolving and a recent development for this language model has been the integration to the Bing search engine.¹⁴ This integration allows ChatGPT to be used with portable document

format (PDF) documents for data extraction. This article will describe how Bing AI and MICROSOFT EDGE can be used to perform data extraction and discuss the future of this technology.

2 | METHOD FOR USING BING AI AND MICROSOFT EDGE FOR DATA EXTRACTION IN A SYSTEMATIC REVIEW

2.1 | Aim

This paper will aim to describe a method of a single reviewer carrying out independent data extraction first and then using Bing AI as a second reviewer to verify/enhance the accuracy of data extraction.

2.2 | Requirements

At this point, using Bing AI as a primary mode of data extraction is strongly ill-advised without further examination and testing of its accuracy in multiple domains and scenarios. For the purpose of this article, the instructions will be based around the use of Microsoft Edge (Version 119.0.2151.44 (Official build) (64-bit)) on a personal computer running Windows 11 (09/06/2023). For this demonstration, a recent systematic review completed within the team, looking at the effectiveness of diagnostic stewardship, will be used as example data. The use of a PDF version of all included studies is recommended for this process. Bing AI can data extract from website versions of papers, however, this may lead to possible additional errors due to non-relevant information being on the webpage. Furthermore, it is important to note that this method does not work well for conference proceedings as these documents typically have more than one article within the PDF making it difficult for the system to identify the specific paper.

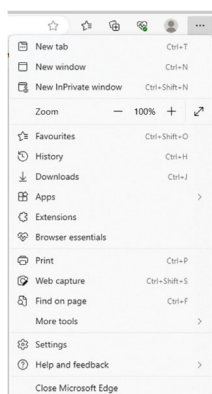
2.3 | Stage I: Independent data extraction (reviewer one)

1. For this instructional paper, it is assumed that you already have fundamental knowledge regarding data extraction methodology. For further guidance on how to undertake data extraction please see the Cochrane Handbook Chapter 5.¹⁵ The first stage is to undertake data extraction independently by a single reviewer, using your chosen pre-piloted data extraction form. In cases where it is suitable, we recommend using Excel for a seamless process of copying and pasting results, making it easier to compare them at later stages within this method. It is essential to carry out this step first to avoid any type of bias which may be introduced by using the AI system first. If using a table for data extraction it is important to place the data items within the columns and place the extracted data within the rows, as this will make it easier for the matching of Bing AI outputs. See Table 1 for example.

2.4 | Stage II: Bing automated data extraction (verification of reviewer one)

For the second stage, users should ensure that the most up-to-date version of Microsoft Edge is installed on your given device, and you are logged into a Microsoft account within the browser.

1. To check you are running the most up-to-date version of Microsoft Edge, open the application and click on the three dots in the top righthand corner. Select “Settings” on the drop-down menu and then click on “About Microsoft Edge” on the left side of the screen. If a message appears stating “Microsoft Edge is up to date” then no further action is required, if the message does not appear updates will need to be downloaded.



(Microsoft Edge, 2023).¹⁷

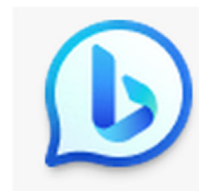
2. To login to use Bing AI you will need to Click on the profile image located in the browser taskbar of Microsoft Edge



(Microsoft Edge, 2023).¹⁷

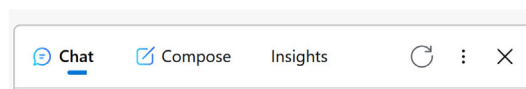
3. From the available options, select the account you wish to sign in to. Alternatively, you can choose “Add new account” if you want to add a new account. To synchronise your data, click on “Sign in to sync data.” If you see your Microsoft email address, it means you are already signed in. Certain organisations might oversee the Bing Chat sidebar and restrict the user from enabling it, even if they have set up and signed into a personal account. In this case it may be preferable to conduct the work on your own personal device rather than a work device.

4. After signing in the next step is to click on the “Bing AI” symbol in the top right corner.



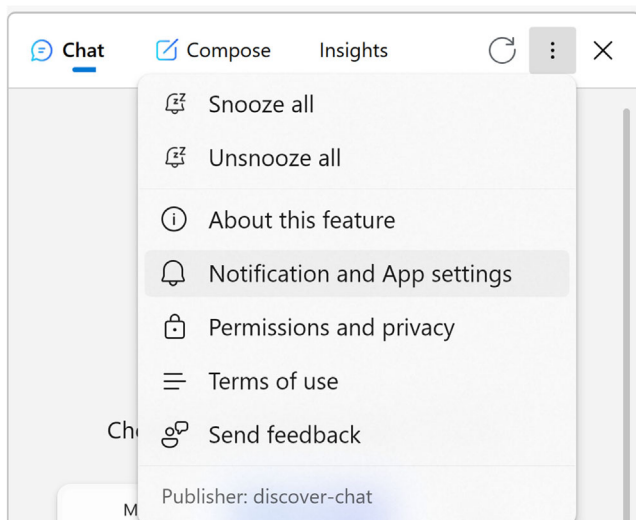
(Microsoft Edge, 2023).¹⁷

5. By clicking on the symbol you will be presented with the three options, see below. If the options below do not appear in the sidebar, click on the three horizontal dots in the top-right hand corner of the browser, go to “Settings,” click “Reset Settings,” and then in the Reset settings window click “Restore setting to their default values” to reset the browser. If at subsequent stages in the process, “Bing AI” resets to the “Discover” and the three options disappear, users may need to “Reset Settings” again.



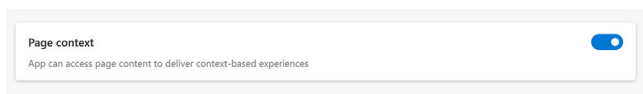
(Microsoft Edge, 2023).¹⁷

6. The next step is to click on the vertical three dots in the sidebar and then select the “Notification and App settings.”



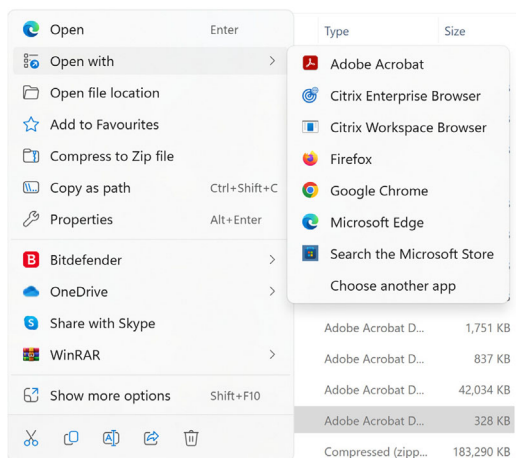
(Microsoft Edge, 2023).¹⁷

7. In the settings, you need to allow the app to gain access to the webpage by ensuring that the setting for “Page context” option is selected.



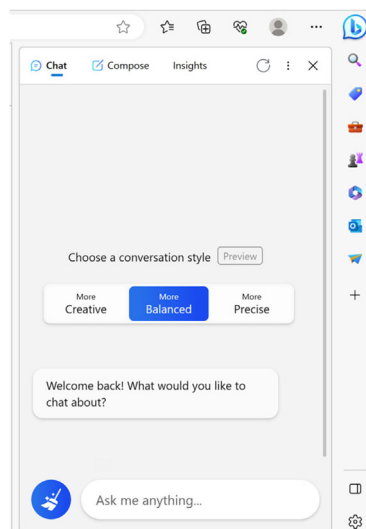
(Microsoft Edge, 2023).¹⁷

8. After this is selected, you will first need to open up the PDF of the paper for data extraction. This can be performed by simply dragging the PDF document into the web browser or alternatively right click on the PDF document and select “Open with” and then “Microsoft Edge.”



(Microsoft Edge, 2023).¹⁸

9. When the PDF document is opened in Microsoft Edge click back on the Bing AI icon and select “Chat.” We recommend using a “Balanced” setting initially but this may need to be adjusted to “Precise” for certain domains.



(Microsoft Edge, 2023).¹⁷

10. The next stage is to enter the command for data retrieval within the “Ask me anything...” field. This command will need to ask the system to identify your data items which you included in your data extraction form. At the end of your data items, you need to indicate for the system to place this information in a table. The request below was given without any prior training or development¹⁹ and was able to accurately retrieve the given data items for the example study.

“What is the aim, country, town, type of study, population, sample size, clinical setting of this study. Can you place the information in a horizontal table.”

As can be seen within Table 2, Bing AI was able to accurately retrieve the data items. However, it is important to note Bing AI will not be able to easily generate coded data items, for example, as you can see within type of clinical setting column. If the system does not produce the appropriate output you require, it is important to adjust the enquiry and repeat until you feel that the system is producing a reliable and consistent output for that data item. For example, if the term “sample size” is used but you are particularly interested in the number of blood samples taken rather than number of participants in the study, you may need to refine your question to specifically refer to the test sample type. As an example presented in Table 2, the number retrieved regarding sample size was based upon individuals in the study rather than samples taken.

TABLE 1 Example layout of a data extraction table.

Study name	Aims of study	Country	Town/city	Study type	Population	Sample size at start of study	Clinical setting
Fleming et al ¹⁶	The aim of the study was to assess the influence of computerised clinical decision support (CCDS) on the appropriateness of C. difficile testing.	United States of America	Maryland	Before and after study	Hospitalised patients undergoing C. difficile testing	280 patients	Secondary care medicine unit

TABLE 2 Data extraction single reviewer versus Bing AI data extraction.

Data extraction method	Aim	Country	Town	Type of study	Population	Sample size	Clinical setting
Single reviewer data extraction	The aim of the study was to assess the influence of computerised clinical decision support (CCDS) on the appropriateness of C. difficile testing.	United States of America	Baltimore	Before and after study	Hospitalised patients undergoing C. difficile testing	280 patients	Secondary care medicine unit
Bing AI data extraction	Evaluate the impact of computerised clinical decision support (CCDS) on the appropriateness of C. difficile testing	The United States of America	Baltimore, MD	Quasi-experimental study	Consecutive hospitalised patients undergoing C. difficile testing from 2/19/16 to 3/19/16	280 patients in pre-intervention group and 167 patients in post-intervention group	Academic hospital

If needed, you can ask for additional detail on any given item by requesting further information on the given data item, for example “tell me more about the aims of the study.” After the system has completed generating the table, if your data extraction table has been structured in a horizontal format (as indicated above), the information should be easily copied and pasted directly into Excel. To continue data extraction, select the next PDF document for data extraction and repeat from 2.8 onwards.

3 | DISCUSSION

3.1 | Strengths and weaknesses of the approach

From our initial assessments, this technique may provide the opportunity to add an additional verification process for extracting simple data items for either novice reviewers or where there are limited resources. Currently this system is scalable and has no known capacity restrictions. Furthermore, there does not seem to be any plans

to make this functionality within Microsoft Edge a paid service. Therefore, outside the generation of the command and opening up and running the enquiry for each paper, there are minimal staffing costs compared to traditional data extraction methods. At present, Bing AI does not provide any methods to undertake more complicated data retrieval processes, such as assessment of bias. It is also important to be aware that for more complicated data extraction requests, the accuracy of retrieval may vary. Furthermore, this system may also provide substantial variation in accuracy of data retrieval in more complex domains. Currently Bing AI does not offer the ability to tweak the model to optimise recall outside the global conversational style categories.

3.2 | Future research

This technique is in its infancy, however, as with all systematic review methodological processes it is essential to ensure transparency and repeatability.²⁰ Moving forward, it will be important to create a standardised reporting

mechanism, which details the exact role the AI system took within the data extraction process. The PRISMA 2020 currently only makes specific reference to automated tools regarding study selection.²¹ However, the PRISMA reporting standard does mention specific principles which may be transferable to data extraction. In the context of using AI for data extraction these would be, first to indicate the exact method of how the AI was used for data extraction (e.g., second reviewer, verification only and first reviewer was independent of AI results). Second, to indicate the version of the technology, as this technology is constantly evolving it would be appropriate to also indicate the date of use. Third, if any additional processes were used regarding “training the system” prior to undertaking the data extraction process. This will be a description of any type of pre-set questions or enquiries which were given to the system prior to requesting the data extraction. Fourth, where this system is used, any validation process should be reported with levels of agreement achieved. Finally, as this is an enquiry-based system the exact enquiry used to request information should be given in full.

As applications develop, it would be helpful if the system could provide greater transparency in exactly what information from the document is being used to generate each individual data item. Within similar data extraction systems such as RobotReviewer, this allows clear tracking of what information has been used to generate each individual data item,²² which helps to reassure that the data items identified are valid. Despite this transparency, recent studies in this area have indicated that there is still substantial scepticism regarding the adoption of automated data extraction systems.²³ Therefore, improved transparency and further validation of Bing AI is required.

3.3 | Future research

Future research on using Bing AI for data extraction (as described in this paper) should explore the recall, precision, efficiency, reliability (extraction time), completeness, and user satisfaction compared to traditional double verified data extraction. Given the characteristics of the generative AI employed in Bing, it becomes crucial for the predictive model to consistently generate results. Consequently, it will be essential to include additional evaluations of repeatability as a component of future research efforts. If feasible, a separate, independent, double-verified gold standard data extraction dataset should be generated for comparison of the two methods.²⁴ An independent assessor should compare both methods to the gold standard with clear and transparent criteria regarding both positive data extraction

and completeness. Future research may want to explore possible moderating factors of experience of reviewers, subject domain, complexity of data items and level of Bing AI training/question development prior to data extraction.

4 | CONCLUSION

Bing AI may be used to enhance a single reviewer data extraction method or facilitate in helping a novice single reviewer to improve their data extraction methods. However, without verification it should not be seen as a replacement to already established and validated double independent data extraction methods. It is important that future use and research in this area uses transparent and standardised methods of reporting of use. Future research should focus on assessing key outcomes and possible moderating factors.

AUTHOR CONTRIBUTIONS

James Edward Hill: Conceptualization; methodology; software; validation; writing – review and editing; writing – original draft; project administration; visualization. **Catherine Harris:** Methodology; software; writing – review and editing; writing – original draft; validation; visualization. **Andrew Clegg:** Writing – review and editing; writing – original draft; visualization; validation; methodology.

ACKNOWLEDGEMENTS

We would like to acknowledge the National Institute for Health and Care Research Applied Research Collaboration North West Coast. As this research was partly-funded by the National Institute for Health and Care Research Applied Research Collaboration North West Coast (NIHR ARC NWC). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

FUNDING INFORMATION

This research was partly-funded by the National Institute for Health and Care Research Applied Research Collaboration North West Coast (NIHR ARC NWC). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

James Edward Hill  <https://orcid.org/0000-0003-1430-6927>

REFERENCES

- Jonnalagadda SR, Goyal P, Huffman MD. Automating data extraction in systematic reviews: a systematic review. *Syst Rev*. 2015;4:78. doi:10.1186/s13643-015-0066-7
- Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open*. 2017;7(2):e012545. doi:10.1136/bmjopen-2016-012545
- Sampson M, Shojania KG, Garrity C, Horsley T, Ocampo M, Moher D. Systematic reviews can be produced and published faster. *J Clin Epidemiol*. 2008;61(6):531-536. doi:10.1016/j.jclinepi.2008.02.004
- Tawfik GM, Dila KAS, Mohamed MYF, et al. A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Trop Med Health*. 2019;47:46. doi:10.1186/s41182-019-0165-6
- Götzsche PC, Hróbjartsson A, Maric K, Tendal B. Data extraction errors in meta-analyses that use standardized mean differences. *JAMA*. 2007;298(4):430-437. doi:10.1001/jama.298.4.430
- Mathes T, Klaffen P, Pieper D. Frequency of data extraction errors and methods to increase data extraction quality: a methodological review. *BMC Med Res Methodol*. 2017;17(1):152. doi:10.1186/s12874-017-0431-4
- Marshall IJ, Wallace BC. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Syst Rev*. 2019;8(1):163. doi:10.1186/s13643-019-1074-9
- Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. *J Am Med Inform Assoc*. 2011;18(5):544-551. doi:10.1136/amiajnl-2011-000464
- Mutinda FW, Liew K, Yada S, Wakamiya S, Aramaki E. Automatic data extraction to support meta-analysis statistical analysis: a case study on breast cancer. *BMC Med Inform Decis Mak*. 2022;22:158. doi:10.1186/s12911-022-01897-4
- van de Schoot R, de Bruin J, Schram R, et al. An open source machine learning framework for efficient and transparent systematic reviews. *Nat Mach Intell*. 2021;3(2):125-133. doi:10.1038/s42256-020-00287-7
- Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare (Basel)*. 2023;11(6):1-20. doi:10.3390/healthcare11060887
- Qureshi R, Shaughnessy D, Gill KAR, Robinson KA, Li T, Agai E. Are ChatGPT and large language models “the answer” to bringing us closer to systematic review automation? *Syst Rev*. 2023;12(1):72. doi:10.1186/s13643-023-02243-z
- Homolak J. Opportunities and risks of ChatGPT in medicine, science, and academic publishing: a modern promethean dilemma. *Croat Med J*. 2023;64(1):1-3. doi:10.3325/cmj.2023.64.1
- Mehdi Y. Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web. <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>
- Li T, Higgins JPT, Deeks JJ. Chapter 5: Collecting data. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.4* (updated August 2023). Cochrane; 2023. Available from www.training.cochrane.org/handbook
- Friedland AE, Brown S, Glick DR, Lusby MC, Lemkin D, Leekha S. Use of computerized clinical decision support for diagnostic stewardship in *Clostridioides difficile* testing: an academic hospital quasi-experimental study. *J Gen Intern Med*. 2019;34(1):31-32. doi:10.1007/s11606-018-4659-4
- Microsoft. Microsoft Edge (Version 119.0.2151.44 (Official build) (64-bit)). 2023.
- Microsoft. Microsoft PC Windows Feature Experience Pack 1000.22677.1000; 2023.
- van de Sande D, Van Genderen ME, Smit JM, et al. Developing, implementing and governing artificial intelligence in medicine: a step-by-step approach to prevent an artificial intelligence winter. *BMJ Health Care Inform*. 2022;29(1):1-8. doi:10.1136/bmjhci-2021-100,495
- Haddaway NR, Woodcock P, Macura B, Collins A. Making literature reviews more reliable through application of lessons from systematic reviews. *Conserv Biol*. 2015;29(6):1596-1605. doi:10.1111/cobi.12541
- Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*. 2021;372:n160. doi:10.1136/bmj.n160
- Soboczenski F, Trikalinos TA, Kuiper J, Bias RG, Wallace BC, Marshall IJ. Machine learning to help researchers evaluate biases in clinical trials: a prospective, randomized user study. *BMC Med Inform Decis Mak*. 2019;19(1):96. doi:10.1186/s12911-019-0814-z
- Jardim PSJ, Rose CJ, Ames HM, Echavez JFM, Van de Velde S, Muller AE. Automating risk of bias assessment in systematic reviews: a real-time mixed methods comparison of human researchers to a machine learning system. *BMC Med Res Methodol*. 2022;22(1):167. doi:10.1186/s12874-022-01649-y
- Walker VR, Schmitt CP, Wolfe MS, et al. Evaluation of a semi-automated data extraction tool for public health literature-based reviews: Dextr. *Environ Int*. 2022;159:107025. doi:10.1016/j.envint.2021.107025

How to cite this article: Hill JE, Harris C, Clegg A. Methods for using Bing's AI-powered search engine for data extraction for a systematic review. *Res Syn Meth*. 2023;1-7. doi:10.1002/jrsm.1689