Applying the Non-adoption, Abandonment, Scale-up, Spread and Sustainability (NASSS) Framework to evaluate automated evidence synthesis in health behaviour change

Commentary

Peter Branney[1]* , Marta M. Marques[2] & Emma Norris[3]

*Corresponding author: p.branney@bradford.ac.uk

[1] Department of Psychology, School of Social Sciences, Faculty of Management, Law & Social Sciences, University of Bradford, UK; https://orcid.org/0000-0002-2084-461X

[2] Comprehensive Health Research Centre (CHRC), NOVA Medical School, Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Portugal; https://orcid.org/0000-0002-4797-9557

[3] Department of Health Sciences; Brunel University London, UK; https://orcid.org/0000-0002-9957-4025

Abstract (179/200 words)

Automated tools to speed up the process of evidence synthesis are increasingly apparent within health behaviour research, however, frameworks to evaluate the development and implementation of such tools are not routinely used. This commentary explores the potential of the Non-adoption, Abandonment, Scale-up, Spread and Sustainability framework (NASSS; Greenhalgh et al., 2017) for supporting automated evidence synthesis in health behaviour change by applying it to the ongoing Human Behaviour-Change Project, which aims to revolutionise evidence synthesis within behaviour change intervention research. To increase the relevance of NASSS for health behaviour change, we recommend i) terminology changes ('condition' to 'behaviour' and 'patient' to 'end user') and ii) a that it is used prospectively so that complexities can be addressed iteratively. We draw three conclusions about i) the need to specify the organisations that will use the technology, ii) identifying what to do if interdependencies fail and iii) even though we have focused on automated evidence synthesis, NASSS would arguably be beneficial for technology developments in health behaviour change more generally, particularly for invention development (e.g. for a behaviour change app).

Keywords: NASSS framework, automated evidence synthesis, health behaviour change, Human Behaviour-Change Project

Disclosure of Interest: Peter Branney reports to conflict of interest. Emma Norris and Marta Marques were Research Associates on the Human Behaviour-Change Project (HBCP) from 2017 and 2020 and 2017 to 2018, respectively. Both authors remain as collaborators in the project and Associates of UCL's Centre for Behaviour Change.

Data availability statement: Application of the NASSS framework to HBCP in this commentary is based on the available published evidence on the HBCP via published papers, information on the project's Open Science Framework pages **(**https://osf.io/efp4x/) and the project' website (https://www.humanbehaviourchange.org/).

Geolocation: United Kingdom

Contributor Statement: Emma Norris: Conceptualization, Writing - original draft, and Writing - review & editing. Marta Marques: Writing - review & editing. Peter E. Branney: Conceptualization, Writing - original draft, and Writing - review & editing.

Applying the Non-adoption, Abandonment, Scale-up, Spread and Sustainability (NASSS) Framework to evaluate automated evidence synthesis in health behaviour change

The aim of this commentary is to outline the potential of the Non-adoption, Abandonment, Scale-up, Spread and Sustainability framework (NASSS; Greenhalgh et al., 2017) to support and evaluate the development of automated evidence synthesis tools in health behaviour change. Evidence synthesis methodologies, such as systematic reviews and meta-analyses, are essential to interpret complex bodies of knowledge in any given domain, such as health behaviour change (Michie et al., 2017). However, evidence synthesis outputs are often written for an academic audience and hence may be of limited use in commissioning, implementing and evaluating health services (Glasziou et al., 2014). Additionally, evidence synthesis research is highly resource-intensive, with reviews often out-of-date by the time of completion (Elliott et al., 2014). Automated tools are being applied to speed up the process of evidence synthesis in the behavioural and health sciences (Marshall et al., 2020), such as automated data extraction (Jonnalagadda et al., 2015) and crowd-sourcing of article screening (Noel-Storr et al., 2021). Living evidence reviews typically use automated evidence synthesis (Millard et al., 2019; Thomas et al., 2017), to support the process of updating the review as new papers become available (e.g., Cochrane Collaboration, 2019; Simons et al., 2021)

The Human Behaviour-Change Project

The Human Behaviour-Change Project (HBCP) applies artificial intelligence to automatically synthesise evidence on behaviour change interventions from published intervention reports (Mac Aonghusa & Michie, 2021), to provide solutions to the 'big question' of behaviour change: 'What works, compared to what, for what behaviours, how well, for how long, with whom, in what setting and why?' (Michie, West, et al., 2020). Data from intervention evaluation reports of randomised controlled trials is extracted into the HBCP knowledge system using the structure of a Behaviour Change Intervention Ontology (BCIO; Michie, West, et al., 2020). The BCIO, as consistent with other ontologies, provides a set of (1) unique, unambiguous entities (such as objects, attributes and processes), (2) labels and definitions for these entities and (3) specified relationships between these entities (Arp et al., 2015), within the specific context of behaviour change interventions.

The HBCP knowledge system has to-date been piloted using smoking cessation behaviour change intervention papers included within Cochrane reviews (Bonin et al., 2020) and physical activity intervention papers (Michie et al., 2020). HBCP aims to identify where the most effective interventions and robust evidence exists ('known knowns') and gaps in research to be filled ('known unknowns' (Hagger et al., 2020)). When launched, the online HBCP web-system aims to have interfaces tailored to different stakeholders, e.g public, practitioners and policy-makers (Michie et al. 2017). In relation to Technology Readiness Levels (EARTO, 2014), HBCP can be seen as currently at Level 3 (proof of concept) or 4 (validation of prototype in laboratory), with user testing underway in researcher, practitioner and public groups.

The Non-adoption, Abandonment, Scale-up, Spread and Sustainability (NASSS) framework

NASSS is a framework for understanding the non-adoption, abandonment, scale-up, spread and sustainability of technology within a complex healthcare system composed of many interacting entities (Abimbola et al., 2019; Greenhalgh et al., 2018). The NASSS framework has seven domains, each with a set of questions to evaluate the health technology (see Table 1). According to each domain, a system is evaluated either retrospectively or prospectively as 'simple' (straightforward, predictable, and with few components), 'complicated' (multiple interacting components or issues) or 'complex' (dynamic, not easily disaggregated into constituent components, and unpredictable; Greenhalgh et al., 2017). Example technologies evaluated using NASSS include electronic decision support in cardiovascular treatment (Abimbola et al., 2019) and internet-delivered CBT for insomnia (Kadesjö Banck & Bernhardsson, 2020). Our application of the NASSS framework to HBCP in this commentary is based on the available published evidence on the HBCP via papers, information on the project's Open Science Framework pages (West et al., 2016) and the project' website[1].

Using NASSS to inform automated evidence synthesis projects in health behaviour change

Our application of NASSS to HBCP is illustrated in Table 1. From this application, we propose two recommendations to increase the relevance of NASSS to health behaviour change. First, terminology used within NASSS could be adjusted to better relate to health behaviour change contexts. In Domains 1, changing the 'condition' to 'behaviour' would allow clearer specification of outcome

[1] https://www.humanbehaviourchange.org/

behaviours within interventions, and also reflect that behaviour's influence multiple conditions and/or illnesses. HBCP, for example, has been initially piloted on smoking cessation interventions, not for a specific health condition. In Domain 4, changing 'patient' to 'end user' would better reflect that end users of behaviour change interventions are not necessarily patients. These adjustments would allow NASSS to consider the benefit of health technologies for behaviour change and avoid a narrow focus on ill health.

Second, we recommend that a full, prospective NASSS evaluation be performed at the early stages of project conceptualisation, such as at grant writing stage, to ensure implementation complexities are identified and addressed iteratively from the project's start. Although the NASSS was developed for either prospective or retrospective application (Greenhalgh et al., 2017) and it is beneficial to retrospectively consider what has led to the success or failure of different technologies, we argue that the utility of NASSS assessment is in prospectively considering potential challenges in technology development to ensure scale-up, spread and sustainability. In other related applications of NASSS, Shaw et al. (2019) used it as a framework for exploring machine learning in healthcare generally but our suggestion is that it should be used on specific technologies, such as HBCP. Meinert et al., (2020) suggested iterative use of NASSS for an app to reduce social isolation during COVID-19 social distancing measures, similar to our suggestion but providing little detail. Similar to the RAG rating system in project management, where a technology is assessed as either complicated or complex, we would need to find ways of making them simpler. Where this is impractical for some domains, it is nevertheless important for the technology's success to do so across as many domains as possible.

We applied the NASSS to HBCP as a case study (Table 1) and drew three main conclusions. First, specifying the organisations (Domain 5) that would use the technology would have minimised the questions we could not answer (the 'unknowns'). Second, as essential interdependencies with other systems make evidence synthesis technologies complex, it is important to plan what will happen if interdependencies fail. Last, HBCP is an "ambitious project" (Michie et al., 2017, p. 11) and we would therefore not expect all or most NASSS domains to be assessed as 'simple' but using NASSS is nevertheless an important exercise to help us think about the uptake of automated evidence synthesis.

Conclusion

In this commentary, we have argued that NASSS should be used prospectively to enhance the development of sustainable automated evidence synthesis technologies for health behaviour change. NASSS would also arguably be beneficial for technology developments across health behaviour change more generally, such as in intervention development. We hope that with the example provided in this commentary other health behaviour change researchers and interventionists can use this as a basis to implement NASSS in their technology projects or products.

Table 1. Application of the Non-adoption, Abandonment, Scale-up, Spread and Sustainability (NASSS) Framework to the Human Behaviour-Change Project (HBCP)

Domain name	Questions	Our assessment of HBCP for each domain
1: The condition	1A: What is the nature of the health behaviour?	*In this domain, we have reframed 'condition or illness' as 'health behaviour' to better reflect the behavioural outcomes that are of interest within the HBCP system.* HBCP has initially been piloted using data from smoking cessation intervention papers, due to the behaviour's relatively large and well-specified evidence base (Michie et al., 2017). HBCP is also experimenting with papers on physical activity as a second behaviour (Michie, Thomas, et al., 2020). These health behaviours, as well as any behaviour, are complex to define and measure, whilst being reported in vastly varying ways within papers (Michie et al., 2017). **Complicated:** Smoking and physical activity as health behaviours currently being piloted within the HBCP knowledge system are reported in vastly varying ways within papers, making the data complicated to characterise.
	1B. What are the relevant sociocultural factors and comorbidities?	*Sociocultural factors:* Socio-cultural factors of participant ethnicity, sexual orientation and religion are extracted within the Intervention Population Ontology (as yet unpublished; Michie et al., 2020) and geographical location and urban/rural setting are recorded within the Intervention Setting Ontology (Norris et al., 2021)(Norris et al., 2021). *Comorbidities:* 'Relevant comorbidities' within HBCP can be seen in conditions related to the target behaviour being addressed within an intervention, e.g lung cancer in the case of the smoking behaviour change interventions. Effects of interventions on longer-term comorbidities are not often explored within behaviour change studies (Bonin et al., 2020), usually focusing instead on shorter-term behavioural or antecedent (e.g attitudes) outcomes. However 'relevant comorbidities' within HBCP could also be seen in potential intentional or unintentional spillover effects onto other related behaviours (Benjamin-Chung et al., 2018). For example, reduction of smoking may be implicated in positive spillover effects of decreased alcohol consumption (e.g., Beard et al., 2017), or other behaviour in related individuals e.g children or spouses (e.g., Fletcher & Marksteiner, 2017). **Complicated:** Participant demographic data is routinely reported within behaviour change intervention reports (Bonin et al., 2020). HBCP and its ontology can provide a clear framework for extracting socio-cultural information. However, effects of behavioural interventions on health conditions and spillover effects into other health behaviours and conditions are not often explored or reported in behaviour change

		interventions. If such outcomes are not assessed or reported in interventions, they cannot be incorporated within an evidence synthesis system.
2. The technology	2A. What are the key features of the technology?	Data across papers is encapsulated and presented within HBCP according to Behaviour Change Intervention Ontology entities (Michie, Thomas, et al., 2020). Published papers are identified using Microsoft Academic Graph (Shemilt et al., 2021) and fed into the bespoke HBCP knowledge system. Algorithms to identify and interpret information coded with ontology entities are being developed (Bonin et al., 2020). Users will interrogate data from the system via a free website interface. **Complex:** Some interdependencies to external knowledge systems (such as to the now defunct Microsoft Academic Graph) are used within HBCP. Bespoke algorithms are being developed to identify and interpret information from papers coded using the Behaviour Change Intervention Ontology.Bespoke user interface being developed.
	2B. What kind of knowledge does the technology bring into play?	The system currently incorporates data from smoking behaviour change intervention reports included within Cochrane reviews (Michie et al., 2017). Integration of wider grey literature and app data could potentially be integrated in future developments to the system to include wider sources of data. **Complex:** Scientific papers follow somewhat common structure but extracting diverse data from text, figures and tables is highly complex (Bonin et al., 2020; Mac Aonghusa & Michie, 2021). Integration of wider grey literature and app data would further increase the complexity of data to be handled in the system.

	2C. What knowledge and/or support is required to use the technology?	Users will need to be guided on how to input a search into the HBCP system and how to interpret outputs and recommendations made by it (Veall et al., 2022). Iterative usability testing is ongoing to understand what information is important and how users want to access it.. This usability testing is being informed by the development and iterative refinement of personas representing academics, students, healthcare commissioners, clinicians, policy-makers etc (Michie, Thomas, et al., 2020). **Complicated:** Clear guidance on how to input searches and interpret outputs and recommendations needs to be developed in collaboration with the diverse user groups that are intended to use the HBCP system. Additional ongoing help desk support may also be required.
	2D. What is the technology supply model?	The HBCP system is aimed to be accessible to anyone via a website that is openly available (Michie et al., 2017). **Simple:** Users will access the system via a free openly available weblink, accessible on any device and browser.
	2E. Who owns the intellectual property (IP) generated by the technology?	The code for HBCP is available on GitHub with an Apache 2.0 licence, which allows for international and perpetual re-use and development for commercial gain internationally. The developed system is to be hosted and operated by UCL via an open access online portal. All algorithm scripts, data feeding into the system etc are available via GitHub and Open Science Framework. **Simple:** Inputs and outputs from the project are designed to be Open Access.
3: The value proposition and value chain	3A. What is the developer's business case for the technology (supply-side value)?	HBCP has been publicly funded through research (Wellcome Trust grant number: 201524). It could produce a return on investment by selling it to healthcare commissioners although it has not been tested on this group. A multinational technology company (IBM) was a collaborator on HBCP, so could have had a role in developing its supply-side value. The code for HBCP has an Apache 2.0 licence, which allows for international and perpetual re-use and development for commercial gain internationally. The Android mobile phone operating system is an example of a commercially successful software that uses the Apache 2.0 licence (Android Open Source Project, 2021). This does mean the intellectual property is free for organisations to further develop HBCP and they can retain copyright for any parts they modify. Nevertheless, this licence does also mean there is a risk of competition that might reduce return on investment. **Complex**: In NASSS, a complex for 3A is based on the business case being 'implausible' compared to 'underdeveloped' for complicated. While the initial investment and licensing suggest possibilities for development, there is no business case in the public domain for HBCP, hence we have opted for

		'complex'.
	3B. What is its desirability, efficacy, safety and cost-effectiveness (demand-side value)?	The value is that AI-supported knowledge synthesis may allow decision makers to quickly query the best available evidence. A range of primary care-led commissioning models have been introduced into Western healthcare systems (e.g., Miller et al., 2012) that attempt to shift from political control to the use of local decision making based on "clinically credible and evidence-based measures" (Department of Health, 2010, p. 21). Indeed, commissioners value evidence-based decisions but that this took up considerable managerial and administrative time (Shaw et al., 2013). For example, the early development work involved assessing needs and reviewing existing evidence to develop the service specification took a minimum of one year. HBCP is being tested with a range of stakeholders (Michie et al., 2017; Veall et al., 2022), although there is no information yet published on the efficacy and cost effectiveness for healthcare commissioning. **Complicated:** Our assessment of desirability is based on the needs of healthcare commissioners and as HBCP has initially been developed for smoking cessation, there is therefore significant risk that HBCP would be ineffective for the range of issues they face (and it would be rated as 'complex'). Nevertheless, HBCP is unlikely to be unsafe and much of the desirability, and cost-effectiveness remains unknown which is why we have gone for 'complicated'.

4: The adopter system	4A. What changes in staff roles, practices and identities are implied?	**Unknown:** These are currently unknown. HBCP should reduce the time required to query the best available evidence and therefore allow healthcare commissions to spend greater time on service specification, such as understanding local needs. This could either reduce the time dedicated to commission a new service or allow for more iterations between, for example, evidence review and service specification. Furthermore, HBCP's browser based search-style platform is likely to be familiar to many digitally literate professional staff and therefore require little training. Nevertheless, these are all theoretical at this point as the staff roles, practices and identities are unspecified.
	4B. What is expected of the end user — and is this achievable by and acceptable to them?	In this domain, we have reframed 'patient and/or immediate carer' as 'end-user of the HBCP system'. Within HBCP, end-users are described to include public and patients, practitioners, researchers, policy-makers and healthcare commissioners (Michie, Thomas, et al., 2020). Users will be required to input search strategies of interest to them on the HBCP system website. Users will need to be guided on how to input a search into the HBCP system and how to interpret outputs and recommendations made by it (Veall et al., 2022). **Complicated:** Users will be required to do routine tasks such as entering search entries and interpreting outputs.
	4C. What is assumed about the extended network of end users	In this domain, we have reframed 'extended network of lay carers' as 'end-user of the HBCP system'. Within HBCP, end-users are described to include public and patients, practitioners, researchers, policy-makers and healthcare commissioners (Michie, Thomas, et al., 2020). **Simple:** No requirement of a network of end users to support users of the system will be made.
5: The organisation(s)	5A. What is the organisation's capacity to innovate?	**Unknown**: Answers to the organisation domain questions have been rated as unknown for two reasons. First, as UK researchers we could focus on NHS healthcare commissioning organisations but this is, at the time of writing, changing. The operational plan is that these changes will occur over 12 months (NHS, 2022) and it is therefore challenging to consider these questions of organisations that are coming into being. Second, these questions need asking of each individual organisation involved and, as far as we are aware, there is no publicly available information through which they could be answered. While healthcare commissioning organisations may have similar legal structures within a country, there is the possibility that they may have different responses to each of the questions for this domain. Consequently, the answers to the questions are currently unknown.
	5B. How ready is the organisation for this technology-supported change?	

	5C. How easy will the adoption and funding decision be?	
	5D. What changes will be needed in team interactions and routines?	
	5E. What work is involved in implementation and who will do it?	
6: the wider system	6A. What is the political, economic, regulatory, professional (e.g medicolegal) and sociocultural context for program roll-out?	*Political context:* HBCP is primarily developed in the UK and Ireland. *Regulatory context:* Regulation is not required to develop a tool to synthesise already published evidence on behaviour change interventions. *Professional bodies*: As mentioned in Domain 5, the nature of NHS healthcare commissioning organisations is currently changing (NHS, 2022). As such, it is unclear how the HBCP system can be implemented within this new structure. Formal collaborations with external organisations to support roll-out of the HBCP system within health systems (e.g NHS, UK Health Security Agency, WHO) are not yet apparent. Recruitment of expert stakeholders for ontology review and system development has been supported by national organisations such as British Psychological Society's Division of Health Psychology and international organisations such as European Health Psychology Society. *Socio-cultural context:* A parallel project to HBCP funded by a Wellcome Enrichment Grant is working to assess and improve public trust in AI within public health (Veall et al., 2022). To improve trust in the end system, all papers inputting into decision-making processes will be made available to users to limit unclear, 'black box' conclusions in AI outputs (Michie et al., 2017). Open Software and Scripts for algorithms underlying the system are available via GitHub and Open Science Framework. The presence of more published papers from higher-income countries means that the HBCP system is primarily based on interventions in higher-income contexts. Only English language papers are included in the system. *Inter-organisational working:* The HBCP system has been developed in collaboration with academic institutions (UCL, University of Cambridge and University of Aberdeen) and an international industry

		partner (IBM Research), funded by a Wellcome Trust grant. **Complicated:** Professional and lay stakeholders are not yet publicly committed to the use and implementation of the HBCP system within healthcare systems.
7: Embedding and adaptation over time	7A. How much scope is there for adapting and co-evolving the technology and the service over time?	**Unknown**: As with Domain 5, the answers to these questions are currently unknown. Indeed, as we highlighted that commissioning organisations in the UK NHS are changinging, this highlights two issues for embedding and adapting technology over time. First, these organisations may be open to changes and therefore to identifying and adapting to issues as they emerge. Second, these organisations may be focused on large-scale organisational changes that means all other changes are on hold. Nevertheless, as these organisations are not yet fully formed, the answers to the questions are currently unknown.
	7B. How resilient is the organisation to handling critical events and adapting to unforeseen eventualities?	

Note: NASSS domains and subdomains adapted from Greenhalgh et al. (2017).

Key;

Simple	Green
Complicated	Amber
Complex	Red

Unknown	Grey

References

Abimbola, S., Patel, B., Peiris, D., Patel, A., Harris, M., Usherwood, T., & Greenhalgh, T. (2019). The NASSS framework for ex post theorisation of technology-supported change in healthcare: Worked example of the TORPEDO programme. *BMC Medicine*, *17*(1), 233. https://doi.org/10.1186/s12916-019-1463-x

Android Open Source Project. (2021). *Licenses*. Android Open Source Project. https://source.android.com/setup/start/licenses

Arp, R., Smith, B., & Spear, A. D. (2015). *Building Ontologies with Basic Formal Ontology*. MIT Press.

Beard, E., West, R., Michie, S., & Brown, J. (2017). Association between smoking and alcohol-related behaviours: A time-series analysis of population trends in England: Smoking and alcohol consumption. *Addiction*, *112*(10), 1832–1841. https://doi.org/10.1111/add.13887

Benjamin-Chung, J., Arnold, B. F., Berger, D., Luby, S. P., Miguel, E., Colford Jr, J. M., & Hubbard, A. E. (2018). Spillover effects in epidemiology: Parameters, study designs and methodological considerations. *International Journal of Epidemiology*, *47*(1), 332–347. https://doi.org/10.1093/ije/dyx201

Bonin, F., Gleize, M., Finnerty, A., Moore, C., Jochim, C., Norris, E., Hou, Y., Wright, A. J., Ganguly, D., Hayes, E., Zink, S., Pascale, A., Mac Aonghusa, P., & Michie, S. (2020). HBCP Corpus: A New Resource for the Analysis of Behavioural Change Intervention Reports. *Proceedings of the 12th Language Resources and Evaluation Conference*, 1967–1975. https://aclanthology.org/2020.lrec-1.242

Cochrane Collaboration. (2019). *Guidance for the production and publication of Cochrane living systematic reviews: Cochrane Reviews in living mode*. https://community.cochrane.org/sites/default/files/uploads/inline-files/Transform/201912_LSR_Revised_Guidance.pdf

Elliott, J. H., Turner, T., Clavisi, O., Thomas, J., Higgins, J. P. T., Mavergames, C., & Gruen, R. L. (2014). Living Systematic Reviews: An

Emerging Opportunity to Narrow the Evidence-Practice Gap. *PLoS Medicine*, *11*(2), e1001603. https://doi.org/10.1371/journal.pmed.1001603

Fletcher, J., & Marksteiner, R. (2017). Causal Spousal Health Spillover Effects and Implications for Program Evaluation. *American Economic Journal: Economic Policy*, *9*(4), 144–166. https://doi.org/10.1257/pol.20150573

Glasziou, P. P., Chalmers, I., Green, S., & Michie, S. (2014). Intervention Synthesis: A Missing Link between a Systematic Review and Practical Treatment(s). *PLoS Medicine*, *11*(8), e1001690. https://doi.org/10.1371/journal.pmed.1001690

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. *Journal of Medical Internet Research*, *19*(11), e367. https://doi.org/10.2196/jmir.8775

Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Procter, R., & Shaw, S. (2018). Analysing the role of complexity in explaining the fortunes of technology programmes: Empirical application of the NASSS framework. *BMC Medicine*, *16*(1), 66. https://doi.org/10.1186/s12916-018-1050-6

Hagger, M. S., Moyers, S., McAnally, K., & McKinley, L. E. (2020). Known knowns and known unknowns on behavior change interventions and mechanisms of action*. *Health Psychology Review*, *14*(1), 199–212. https://doi.org/10.1080/17437199.2020.1719184

Jonnalagadda, S. R., Goyal, P., & Huffman, M. D. (2015). Automating data extraction in systematic reviews: A systematic review. *Systematic Reviews*, *4*(1), 78. https://doi.org/10.1186/s13643-015-0066-7

Kadesjö Banck, J., & Bernhardsson, S. (2020). Experiences from implementation of internet-delivered cognitive behaviour therapy for insomnia in psychiatric health care: A qualitative study applying the NASSS framework. *BMC Health Services Research*, *20*(1), 729.

https://doi.org/10.1186/s12913-020-05596-6

Mac Aonghusa, P., & Michie, S. (2021). Artificial Intelligence and Behavioral Science Through the Looking Glass: Challenges for Real-World Application. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, *54*(12), 942–947. https://doi.org/10.1093/abm/kaaa095

Marshall, I. J., Johnson, B. T., Wang, Z., Rajasekaran, S., & Wallace, B. C. (2020). Semi-Automated evidence synthesis in health psychology: Current methods and future prospects. *Health Psychology Review*, *14*(1), 145–158. https://doi.org/10.1080/17437199.2020.1716198

Michie, S., Thomas, J., Johnston, M., Aonghusa, P. M., Shawe-Taylor, J., Kelly, M. P., Deleris, L. A., Finnerty, A. N., Marques, M. M., Norris, E., O'Mara-Eves, A., & West, R. (2017). The Human Behaviour-Change Project: Harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. *Implementation Science*, *12*(1), 121. https://doi.org/10.1186/s13012-017-0641-5

Michie, S., Thomas, J., Mac Aonghusa, P., West, R., Johnston, M., Kelly, M. P., Shawe-Taylor, J., Hastings, J., Bonin, F., & O'Mara-Eves, A. (2020). The Human Behaviour-Change Project: An artificial intelligence system to answer questions about changing behaviour. *Wellcome Open Research*, *5*, 122. https://doi.org/10.12688/wellcomeopenres.15900.1

Michie, S., West, R., Finnerty, A. N., Norris, E., Wright, A. J., Marques, M. M., Johnston, M., Kelly, M. P., Thomas, J., & Hastings, J. (2020). Representation of behaviour change interventions and their evaluation: Development of the Upper Level of the Behaviour Change Intervention Ontology. *Wellcome Open Research*, *5*, 123. https://doi.org/10.12688/wellcomeopenres.15902.1

Millard, T., Synnot, A., Elliott, J., Green, S., McDonald, S., & Turner, T. (2019). Feasibility and acceptability of living systematic reviews: Results from a mixed-methods evaluation. *Systematic Reviews*, *8*(1), 325. https://doi.org/10.1186/s13643-019-1248-5

NHS. (2022). *Roadmap for integrating specialised services within Integrated Care Systems* (No. PAR1440). https://www.england.nhs.uk/wp-content/uploads/2022/05/PAR1440-specialised-commissioning-roadmap-addendum-may-2022.pdf

Noel-Storr, A., Dooley, G., Elliott, J., Steele, E., Shemilt, I., Mavergames, C., Wisniewski, S., McDonald, S., Murano, M., Glanville, J., Foxlee, R., Beecher, D., Ware, J., & Thomas, J. (2021). An evaluation of Cochrane Crowd found that crowdsourcing produced accurate results in identifying randomized trials. *Journal of Clinical Epidemiology*, *133*, 130–139. https://doi.org/10.1016/j.jclinepi.2021.01.006

Norris, E., Wright, A. J., Hastings, J., West, R., Boyt, N., & Michie, S. (2021). Specifying who delivers behaviour change interventions: Development of an Intervention Source Ontology. *Wellcome Open Research*, *6*, 77. https://doi.org/10.12688/wellcomeopenres.16682.1

Shaw, S. E., Smith, J. A., Porter, A., Rosen, R., & Mays, N. (2013). The work of commissioning: A multisite case study of healthcare commissioning in England's NHS. *BMJ Open*, *3*(9), e003341. https://doi.org/10.1136/bmjopen-2013-003341

Shemilt, I., Arno, A., Thomas, J., Lorenc, T., Khouja, C., Raine, G., Sutcliffe, K., Preethy, D., Kwan, I., Wright, K., & Sowden, A. (2021). *Cost-effectiveness of Microsoft Academic Graph with machine learning for automated study identification in a living map of coronavirus disease 2019 (COVID-19) research*. https://doi.org/10.12688/wellcomeopenres.17141.1

Simons, D., Shahab, L., Brown, J., & Perski, O. (2021). The association of smoking status with SARS-CoV-2 infection, hospitalization and mortality from COVID-19: A living rapid evidence review with Bayesian meta-analyses (version 7). *Addiction*, *116*(6), 1319–1368. https://doi.org/10.1111/add.15276

Thomas, J., Noel-Storr, A., Marshall, I., Wallace, B., McDonald, S., Mavergames, C., Glasziou, P., Shemilt, I., Synnot, A., Turner, T., Elliott, J., Agoritsas, T., Hilton, J., Perron, C., Akl, E., Hodder, R., Pestridge, C., Albrecht, L., Horsley, T., … Pearson, L. (2017). Living systematic reviews: 2. Combining human and machine effort. *Journal of Clinical Epidemiology*, *91*, 31–37.

https://doi.org/10.1016/j.jclinepi.2017.08.011

Veall, C., Hou, Y., West, R., Michie, S., Shawe-Taylor, J., Thomas, J., Finnerty, A., Johnston, M., Aonghusa, P. M., & O'Mara-Eves, A. (2022). *Research Enrichment—Public Engagement*. https://osf.io/2qkza/

West, R., Michie, S., Shawe-Taylor, J., Thomas, J., Finnerty, A., Johnston, M., Aonghusa, P. M., O'Mara-Eves, A., Stokes, G., & Norris, E. (2016). *Human Behaviour-Change Project*. https://doi.org/10.17605/OSF.IO/EFP4X