



# Efficiency over thoroughness in laboratory testing decision making in primary care: findings from a realist review

Claire Duddy, BA (Hons), MA, AFHEA<sup>1\*</sup>, Geoff Wong, MA, MBBS, MD (Res), MRCGP, FHEA<sup>1</sup>

<sup>1</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

## Abstract

**Background:** Existing research demonstrates significant variation in test-ordering practice, and growth in the use of laboratory tests in primary care. Reviews of interventions designed to change test-ordering practice report heterogeneity in design and effectiveness. Improving understanding of clinicians' decision making in relation to laboratory testing is an important means of understanding practice patterns and developing theory-informed interventions.

**Aim:** To develop explanations for the underlying causes of patterns of variation and increasing use of laboratory tests in primary care, and make recommendations for future research and intervention design.

**Design & setting:** Realist review of secondary data from primary care.

**Method:** Diverse evidence, including data from qualitative and quantitative studies, was gathered via systematic and iterative searching processes. Data were synthesised according to realist principles to develop explanations accounting for clinicians' decision making in relation to laboratory tests.

**Results:** A total of 145 documents contributed data to the synthesis. Laboratory test ordering can fulfil many roles in primary care. Decisions about tests are incorporated into practice heuristics and tests are deployed as a tool to manage patient interactions. Ordering tests may be easier than not ordering tests in existing systems. Alongside high workloads and limited time to devote to decision making, there is a common perception that laboratory tests are relatively inconsequential interventions. Clinicians prioritise efficiency over thoroughness in decision making about laboratory tests.

**Conclusion:** Interventions to change test-ordering practice can be understood as aiming to preserve efficiency or encourage thoroughness in decision making. Intervention designs and evaluations should consider how testing decisions are made in real-world clinical practice.

\*For correspondence: claire.duddy@phc.ox.ac.uk

Competing interest: See page 9

Received: 06 July 2020

Accepted: 13 August 2020

Published: 03 February 2021

©This article is Open Access: CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

**Author Keywords:** realist review, realist synthesis, clinical laboratory techniques, primary health care, practice patterns, physicians, clinical decision making, general practice

Copyright © 2020, The Authors;  
DOI:10.3399/  
bjgpopen20X101146

## How this fits in

Research on laboratory testing has long demonstrated variation and growth in the use of tests. Existing reviews have identified lists of factors associated with test-ordering behaviour, and mixed evidence of the effectiveness of interventions designed to change testing practice. This realist review presents explanations for clinicians' test-ordering behaviour, illustrating the wide range of influences affecting decision making about laboratory testing, and highlighting the combined effect of high workload and a generalised perception that laboratory tests are relatively trivial and inconsequential interventions. As a result, clinicians often prioritise efficiency and pragmatism over thoroughness

in making decisions about the use of laboratory tests, focusing their limited time and resources elsewhere. Future intervention designs and evaluations should take account of real-world practice by recognising that making changes to wider systems may not change perceptions of laboratory testing, and that encouraging thoroughness in decision making in this area of practice may have unintended consequences elsewhere.

## Introduction

Existing research has long demonstrated growth in the use of laboratory tests in primary care and the existence of variation in test-ordering practice.<sup>1–6</sup> These patterns raise important questions about how much variation in clinical decision making is warranted, and whether increased testing improves health outcomes and represents cost-effective use of scarce resources. In the UK, NHS Improvement estimates expenditure of £2.2 billion annually on pathology services.<sup>7</sup> Primary care makes a significant contribution: in 2006, the Carter Review estimated that 35%–45% of requests for laboratory tests originated in primary care;<sup>8</sup> and in 2014, NHS England estimated that over 50 million electronic reports of laboratory test results were delivered to GPs each year.<sup>9</sup>

Laboratory testing often represents an early step in a clinical pathway, carrying further consequences for downstream activity.<sup>10</sup> Both undertesting and overtesting can result in negative consequences for patients. Undertesting can mean delayed or missed diagnoses, and a lack of monitoring of long-term conditions or medication side effects. However, growth in test use raises concerns about overtesting: some testing may be unnecessary because of the low likelihood of benefitting patients,<sup>11</sup> or may even cause harm; for example, unnecessary testing has the potential to increase patient anxiety, raises the chances of false positive results, and has the potential to provoke ‘cascades’ of further unnecessary investigations and interventions.<sup>10,12,13</sup>

Multiple reviews have attempted to assess the effectiveness of a range of interventions designed to change test-ordering behaviour.<sup>14–23</sup> These reviews report heterogeneity in intervention designs, effectiveness, and the sustainability of changes in practice. Across the reviews, the most frequently measured outcomes relate to test-ordering activity and behaviour. There is an emphasis on reducing test ordering or improving ‘appropriateness’. The latter often refers to assessments of whether or not test-ordering activity fits within existing guidelines (see **Table 1** below for more detail). Observed variation in practice and in the impact of interventions suggests multiple causal mechanisms may underlie test-ordering decisions, and that context may play an important role in determining outcomes. Understanding the causes of observed patterns of testing is an important means of informing the development and evaluation of interventions that aim to change practice. The realist approach adopted here aimed to: produce explanations that account for observed patterns and the influence of context; identify the underlying causal mechanisms that underlie test-ordering practice; and produce recommendations to guide future research.<sup>24</sup>

## Method

A realist review is an interpretive form of evidence synthesis, conducted with the aim of identifying and synthesising relevant and trustworthy data that can be used to develop a better understanding of its subject. Realist analysis can be used to develop explanatory theory (called ‘programme theory’), which takes account of important influences of context and identifies underlying causal mechanisms that produce observed outcomes.<sup>24</sup> This realist review aimed to develop explanations for primary care clinicians’ decision making about laboratory tests, and to make theory-informed recommendations for future research and intervention design.

The methods for this review are described in detail in the published protocol.<sup>25</sup> The review was conducted according to Pawson’s five steps,<sup>26</sup> outlined briefly in **Table 2** below and in detail in supplementary **Table S1**. Review processes adhered to the RAMESES quality<sup>27</sup> and reporting<sup>28</sup> standards throughout.

A group of 14 stakeholders, including patients, members of the public, clinicians, a laboratory scientist, and a policymaker were involved from the outset. Their direct knowledge and experience of test-ordering practice shaped the review, contributing to the development of the initial and refined programme theories.

**Table 1** Summary of interventions and outcomes assessed in studies included in existing systematic reviews

Review	Test-ordering outcome(s)	Interventions prioritising efficiency	Interventions prioritising thoroughness			
		Process changes (including computer systems)	Guidelines and/or protocols	Education	Audit and feedback	Financial incentives
Solomon <i>et al</i> 1998 <sup>14</sup>	Reduction in test-ordering volume Reduction in test expenditure	x	x	x	x	x
Main <i>et al</i> 2010 <sup>15</sup>	Changes in test-ordering volume 'Appropriateness' of testing	x				
Smellie 2012 <sup>16</sup>	Reduction in test-ordering volume Reduction in test expenditure 'Appropriateness' of testing	x	x	x	x	x
Cadogan <i>et al</i> 2015 <sup>17</sup>	Reduction in test-ordering volume	x	x	x	x	
Kobewka <i>et al</i> 2015 <sup>18</sup>	Reduction in test-ordering volume	x		x	x	x
Thomas <i>et al</i> 2015 <sup>19</sup>	Reduction in test-ordering volume	x	x		x	
Thomas <i>et al</i> 2016 <sup>20</sup>	Change in test-ordering volume	x		x	x	
Zhelev <i>et al</i> 2016 <sup>21</sup>	Reduction in test-ordering volume Changes in test expenditure 'Appropriateness' of testing Changes in testing patterns	x	x	x	x	x
Delvaux <i>et al</i> 2017 <sup>22</sup>	'Appropriateness' of testing Changes in test expenditure Clinical outcomes	x				
Maillet <i>et al</i> 2018 <sup>23</sup>	Changes in test ordering 'Appropriateness' of testing Workload	x				

## Results

A total of 145 documents contributed data to this review; see **Figure 1**. Most reported research studies ( $n = 123$ ), but grey literature including commentaries ( $n = 20$ ) and theoretical work ( $n = 2$ ) were also included. Documents describing research comprised 60 cross-sectional or survey studies, 27 qualitative studies, 12 narrative reviews, nine systematic reviews, six cohort studies, five decision-analysis studies, two randomised controlled trials, one project evaluation, and one case report. Full details of the included documents are provided in supplementary **Table S5**.

The final programme theory developed from the realist analysis is presented below. This explanatory framework is underpinned by three overarching context–mechanism–outcome configurations (CMOCs) developed during the review, summarised in **Table 3**. The three overarching CMOCs were

**Table 2** Summary of realist review methodology

<b>Step 1</b> Initial programme theory (IPT) development	An IPT is a first attempt to develop an understanding of the research question. To develop the IPT for this review, two scoping searches were run of the literature to identify: (a) existing theoretical perspectives; and (b) common intervention designs in relation to test-ordering practice. Full details of the search strategies are provided in supplementary <b>Table S2</b> . The IPT was further developed via the input of the stakeholder group and is presented in full in supplementary <b>Figure S1</b> .
<b>Step 2</b> Searching for evidence	The main search for evidence was undertaken with the aim of assembling a body of relevant data that could be used to develop and refine the programme theory. A broad range of sources were searched ( $n = 15$ ) to ensure that literature across multiple disciplines was considered. Full details of the main search strategy are provided in supplementary <b>Table S3</b> . Additional documents were identified via supplementary search methods such as citation tracking (snowballing) and via personal contacts and networks. <sup>69,70</sup> Further searches were undertaken later to identify relevant substantive theory to act as a theoretical lens through which to understand the review's overall findings. <sup>71</sup> The search strategies employed are provided in supplementary <b>Table S4</b> .
<b>Step 3</b> Selection and appraisal	Document selection was based on an assessment of relevance (whether or not documents contained data that could be used to develop theoretical explanations ['programme theory']) and rigour (whether data were considered credible in relation to their role in contributing to the theory). <sup>26,72</sup> Included documents provided data on important contexts, mechanisms, and outcomes related to clinician decision making in relation to laboratory test ordering in primary care settings, or provided data related to analogous settings or decisions, or relevant theoretical perspectives. More details on data selection processes are provided in supplementary <b>Table S1</b> .
<b>Step 4</b> Data extraction and organisation	Included documents were read closely and coded in NVivo (version 12 Pro) to organise the data and identify important concepts that could inform the realist analysis. The characteristics of included documents ( $n = 145$ ) are provided in supplementary <b>Table S5</b> .
<b>Step 5</b> Analysis and synthesis	Analysis and synthesis of included data involved the iterative development of realist 'context–mechanism–outcome configurations' (CMOCs). These are theoretical causal explanations describing how important contexts trigger the mechanisms that generate observed outcomes. Members of the stakeholder group provided feedback on the relevance and resonance of the developing theories. CMOC development and refinement continued until the reviewers agreed theoretical saturation was reached. A 'final programme theory' (FPT) was developed after consideration of the full set of CMOCs and drawing on theoretical literature. Full details of the CMOCs developed and illustrative data excerpts are provided in supplementary <b>Tables S6–S8</b> .

developed from the detailed analysis of 52 underpinning CMOCs in all (presented in full in **Tables S6–S8**).

### Competing demands, relative triviality, and decision making

The data included in this review demonstrate that laboratory test ordering fulfils a wide range of roles for clinicians. Test-ordering decisions may be built into clinicians' practice heuristics and fulfil numerous social or strategic roles in managing patient interactions. It is also clear that features of the wider environment — including computer systems, organisational structures, and social or cultural norms — often tend to encourage (or fail to discourage) the use of laboratory tests.

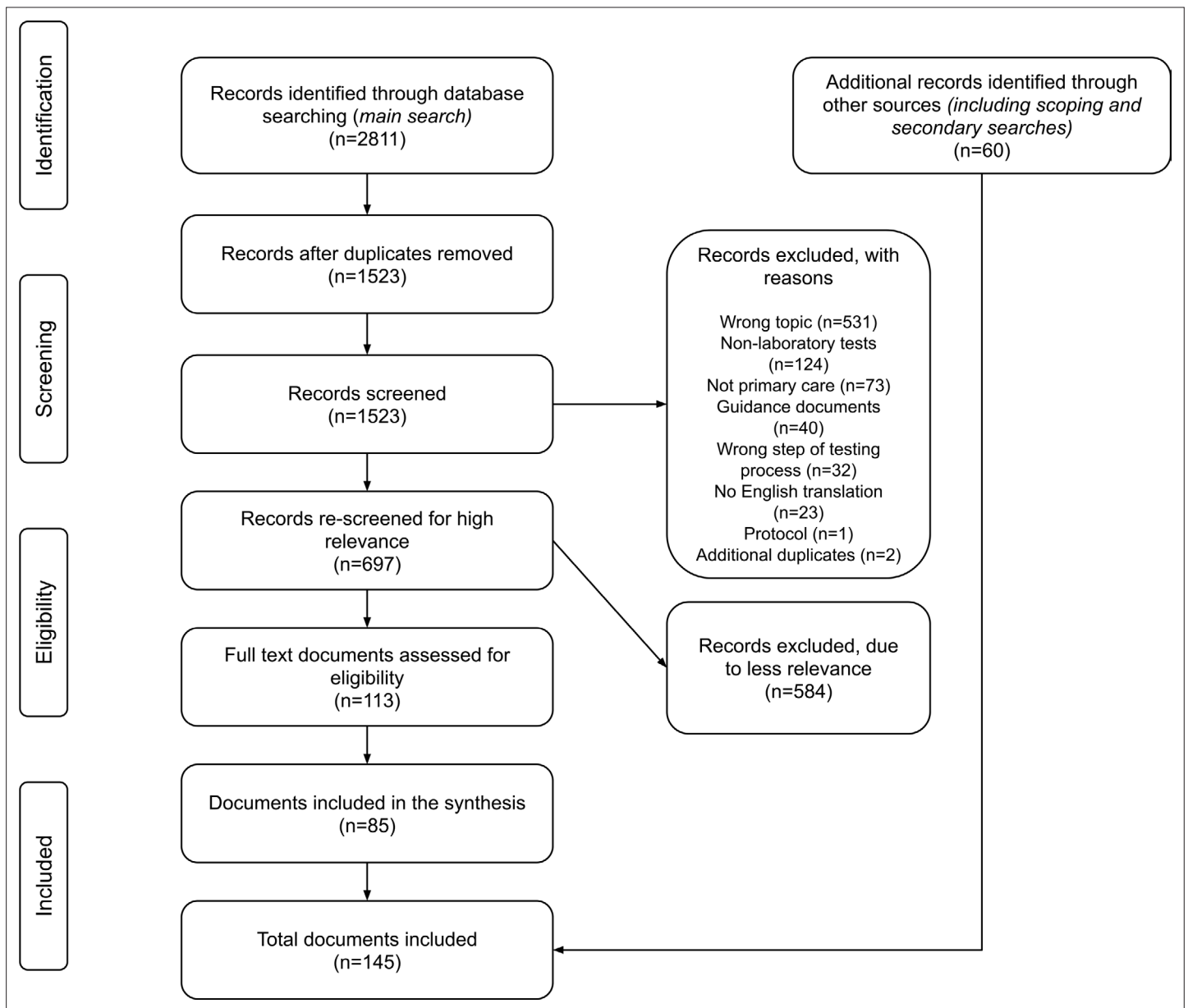
Underpinning these findings are two important overarching contexts: clinicians are juggling heavy workloads and limited time with patients; and laboratory tests are often considered to be relatively trivial and inconsequential interventions.<sup>29–35</sup> Some data even suggest that where there are obvious negative consequences of testing, these may be construed positively by clinicians and patients.<sup>36,37</sup>

Busy clinicians with limited time and attention to devote to many competing tasks must prioritise. Hollnagel's Efficiency–Thoroughness Trade-Off (ETTO) principle provides a framework for understanding this. For example:

*'In their daily activities ... people (and organisations) routinely make a choice between being efficient and being thorough, since it is rarely possible to be both at the same time.'*<sup>138</sup>

On the ETTO spectrum, 'thoroughness' is understood to confer safety, while increases in 'efficiency' sacrifice diligence in favour of saving time and psychological effort. When laboratory tests are considered relatively trivial or inconsequential, clinicians trading off between efficiency and thoroughness may not believe there will be any significant loss of safety in relation to increasing efficiency in decisions about testing. Taking this position permits the application of efficient heuristics and use of tests for social or strategic purposes, and means that clinicians are unlikely to expend effort in working against wider pro-testing systems; see **Figure 2**.

Some data pointed to exceptions: clinicians do resist pressures to test, when they perceive that tests may carry burdens or harms for patients, or when they have adopted professional identities or follow norms associated with more conservative or parsimonious practice (see CMOCs 8a–c, 10c in



**Figure 1** Document screening and selection processes

supplementary **Tables S6–S8**). In such cases, laboratory tests are no longer considered trivial, but are understood to carry real, potentially harmful consequences. However, in these circumstances, clinicians face the same pressures of high workloads and limited time. From the included data, it is unclear if the outcome is more thorough decision making about test ordering (and potentially the prioritisation of efficiency in other areas to compensate) or simply the adoption of heuristics that favour not testing.

In the case of laboratory testing in primary care settings, the dominant mechanism of prioritising efficiency over thoroughness tends to lead to sustained and increasing use of tests.

## Discussion

### Summary

This realist review demonstrates the complexity of laboratory test-ordering practice in primary care. Clinicians use tests to fulfil a variety of roles. Their test-ordering decisions are affected by a wide

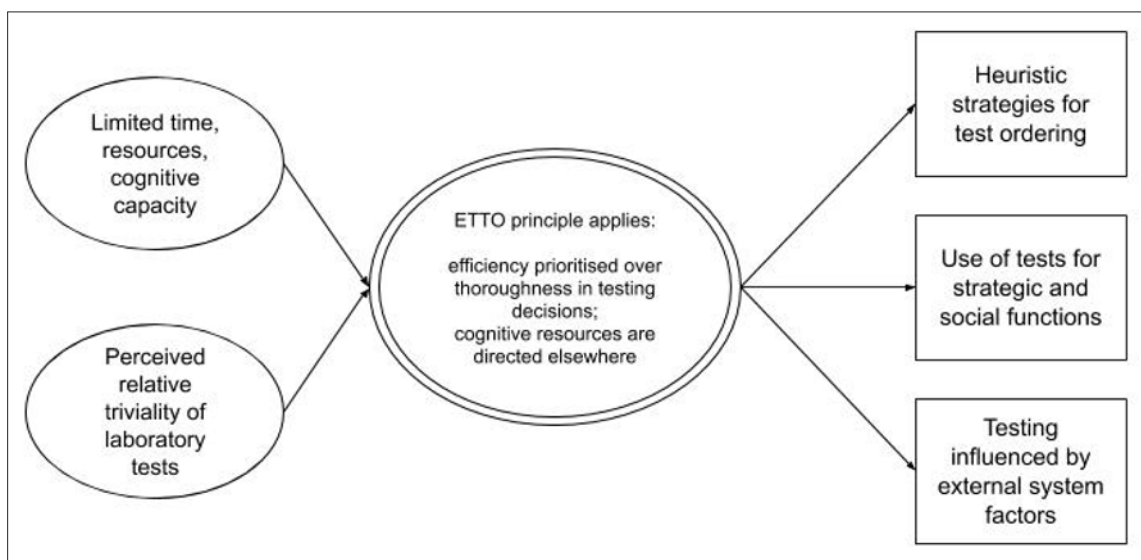
**Table 3** Summary of realist analysis

Overarching CMOCs	Illustrative examples of underpinning CMOCs
When laboratory tests are perceived to be relatively trivial (C1), and cognitive resources are limited (C2), clinicians prioritise efficiency over thoroughness for test-ordering decisions, directing their cognitive resources to other clinical decisions (M) so <b>decisions about testing will be based on heuristics or routines</b> (O).	When clinicians have incomplete technical knowledge about laboratory medicine and/or diagnostic reasoning (C), they rely on 'gist' understanding (M) to develop decision-making heuristics for test ordering (O) [CMOC 1b]. In the presence of diagnostic uncertainty (C), clinicians may apply a heuristic of 'more testing is better' (O1) or 'rule out the worst case' (O2) as they seek to minimise the risk of missing a diagnosis (M) [CMOCs 2a–2b]. When a test or condition is 'in fashion', and there is high awareness among clinicians and/or the public (C), the use of this test may be incorporated into testing heuristics (O) owing to increased awareness ('salience') (M) [CMOC 3g].
When laboratory tests are perceived to be relatively trivial (C1), and cognitive resources are limited (C2), clinicians prioritise efficiency over thoroughness for test-ordering decisions, and direct their cognitive resources to other clinical decisions (M) and so <b>tests may be used to fulfil social and strategic functions</b> (O).	In the presence of diagnostic uncertainty (C), clinicians may demonstrate care (M1), attempt to reassure (M2), or exert control via 'doing something' for their patients (M3) by ordering tests (O) [CMOCs 5a–5c]. When clinicians anticipate a 'difficult' interaction with a patient (C), they may use the offer of a laboratory test (O) as a strategy to help manage the consultation (M) [CMOC 6c]. When clinicians anticipate disagreement with a patient about their proposed management plan (C), they may acquiesce to patient requests or expectations and order tests (O) to avoid having to explain why they are inappropriate (M1) or avoid conflict in the consultation (M2) [CMOCs 7b–7c].
When laboratory tests are perceived to be relatively trivial (C1), and cognitive resources are limited (C2), clinicians will prioritise efficiency over thoroughness in test-ordering decisions, and direct their cognitive resources to other clinical decisions (M) so <b>decisions about testing will be open to wider system influences</b> (O).	When responsibility for patient care is shifted from secondary to primary care (C), clinicians in primary care settings comply with testing expectations and requests (M) received from secondary care and take on responsibility for associated testing (O) [CMOC 9c]. In the absence of disincentives for inappropriate testing (C), clinicians and laboratory managers will not prioritise concerns about under/overtesting (M) and so will not take action to address these problems (O) [CMOC 10b]. When tests are available to order as part of profiles or panels (C), clinicians may try to save time and cognitive energy (M) by ordering full panels instead of individual tests (O) [CMOC 11b].

C = context; M = mechanism; O = outcome

range of contextual factors and generated by many different motivations. Overall, a commonly held perception of laboratory tests as relatively trivial and inconsequential interventions often permits clinicians to prioritise efficiency over thoroughness in test-ordering decisions.

The diversity of roles that laboratory testing can fulfil helps to explain variation in test use as the result of variation in multiple contextual circumstances. Variation can also be understood as resulting from a widespread context of perceived relative triviality, which acts to enable this diversity of motivations for testing and variation in individuals' decision making. However, it is also clear that the



**Figure 2** Final programme theory illustrating overarching context–mechanism–outcome configurations (CMOCs). ETTO = Efficiency–Thoroughness Trade-Off. Figure legend: single line oval = context, double line oval = mechanism, rectangle = outcome.

data suggest that the prioritisation of efficiency in test-ordering practice tends to lead to increased and self-sustaining use of tests. Increasing test use should be seen within the context of shifting norms and expectations in clinical practice, and broader cultural beliefs in the benefits and capabilities of testing and health care.<sup>39</sup>

It should also be understood in the context of primary care workloads,<sup>40,41</sup> and the need for clinicians to find efficient ways of practising with limited resources.<sup>42–44</sup> The review's final programme theory offers a novel framework to understand common patterns of test-ordering behaviour: clinicians with limited time and energy who consider laboratory tests to be relatively trivial interventions are likely to prioritise efficiency over thoroughness in these decisions. They are likely to devote more time and psychological effort to other areas of practice, and are unlikely to expend resources in resisting the multiple societal and system features that tend to incentivise testing.

## Strengths and limitations

This review included a wide range of evidence obtained via systematic searching. The analysis was enhanced by 'borrowing' relevant data from documents focused on other areas of clinical practice and drawing on substantive theory to generate new insight. The involvement of a diverse group of stakeholders helped to shape the project and refine the analysis, ensuring its relevance and resonance for real-world practice and policy.

As in all reviews, findings were limited by the available data. Some plausible explanations for test-ordering behaviour proposed by stakeholders remain unsubstantiated; for example, the question of whether decision making is affected by clinicians' emotional affect, and whether defensive mechanisms may lead to decreased testing where clinicians reason it is best to avoid opening Pandora's box. The available data did not permit determination of which CMOCs are more dominant or explanatory than others. A representative set of CMOCs has been generated offering explanations for observed patterns, but future research may lead to the refinement, confirmation, or refutation of these explanatory theories.

The included literature was variable in quality, and the data underpinning each individual CMOC varies in volume and type. The full details of the contributing data are provided in supplementary **Tables S5–S8** to permit the reader to make judgements about the strength of the evidence underpinning the analysis.

## Comparison with existing literature

Two earlier reviews have collated studies of factors affecting test ordering, similarly highlighting the wide range of influences affecting decision making in this area of practice, including clinicians' experience and attitude toward risk.<sup>45,46</sup> The detailed realist analysis in this review concurs with this image of complexity in test-ordering decision making and extends existing work by focusing on the role of important contexts and the multiple mechanisms that generate clinicians' test-ordering behaviour. The CMOCs presented in supplementary **Tables S6–S8** provide a set of testable theories about clinicians' test-ordering practice that may be refined, confirmed, or refuted by further research. In addition, the final programme theory presented above provides an overarching framework through which to understand the complex picture of testing practice. The application of the ETTO principle as a theoretical lens permits common patterns of variation in test use and growth in the use of tests to be better understood.

The review's findings also have parallels in research conducted in other areas of practice. The ETTO principle has been usefully applied in qualitative studies in primary care, to understand the conduct of medication reviews,<sup>44</sup> and the management of test results,<sup>47</sup> and prescription requests.<sup>43</sup> There are also similarities with findings from other studies of clinical decision making, especially those that employ dual-processing theory to understand how decisions are made.<sup>48,49</sup> The addition of this review's findings further validates the utility of the ETTO principle as a means of understanding the realities of decision making and the prioritisation of workloads in primary care.

## Implications for research

The overall complexity of laboratory test-ordering practice and the perspective provided by the ETTO principle carry important implications for research, and especially for future intervention design and evaluation. The ETTO framework provides a novel way of categorising families of interventions that

aim to influence clinical practice, as aiming either to preserve or increase efficiency, or to encourage thoroughness in decision making. To illustrate this point, the common intervention designs described in the studies included in existing systematic reviews of interventions designed to change test-ordering behaviour are summarised in **Table 1**.

## Efficiency

Intervention designs involving making changes to test-ordering systems (especially order forms) or providing decision support at the point of ordering fall into this category. They aim to adjust clinicians' heuristics or routines by reconfiguring decision-making options, making 'appropriate' testing easy and efficient, or 'inappropriate' testing more difficult. Underlying such interventions is the often implicit theory that clinicians do prioritise efficiency, that is, they base test-ordering decisions on heuristics, informed by their own experience, practice norms, and system constraints, in a similar manner to observed behaviour in relation to guideline adherence.<sup>50</sup>

Reviews report variable effectiveness for interventions that aim to preserve or increase efficiency in primary care settings, although many are reported to result in reductions in testing volumes.<sup>14–21</sup> One study reports that where tests were added to order forms, usage increased,<sup>51</sup> and several studies have demonstrated that adding reminders or alerts can change testing behaviour,<sup>52–59</sup> although 'alert fatigue' may be a problem.<sup>22,60–63</sup>

The wider consequences of relying on decision-making heuristics, and of attempts to impose changes on these routines, which may be 'good enough' and help primary care clinicians manage competing demands and challenging conditions, are unclear and deserve attention. Decision-making heuristics may permit both sound and efficient decision making,<sup>64</sup> or may be vulnerable to errors resulting from cognitive biases.<sup>65</sup> They may also have unintended and unforeseen consequences affecting, for example, clinicians' workloads and relationships with patients over the longer term. Future research, including ethnographic studies of practice and long-term observational studies, could be used to describe in detail, and ultimately assess the reliability of clinicians' testing heuristics, but should prioritise the need to assess patient outcomes associated with testing, rather than the common surrogates of testing volumes or adherence to guidelines (which are frequently consensus based).<sup>66,67</sup>

## Thoroughness

Other interventions seek to focus clinicians' attention on test-ordering practice by adding processes, delivering education, introducing financial incentives, or providing feedback on testing behaviour. As above, reviews report variable effectiveness, although many interventions are successful in changing testing behaviour to some degree.<sup>14–21</sup> As a group, these interventions represent attempts to provoke thoroughness, by providing additional information to factor into decisions or engineering opportunities for reflection on practice. The findings of this review suggest that clinicians' responses may be affected by individual efficiency–thoroughness trade-offs (which may differ from normal practice under trial conditions), informed by perceptions of the relative triviality of laboratory tests, and constrained by workload and competing demands. Where interventions can change one of these important contexts — in practice, perceptions of triviality may be more amenable to change than prevailing conditions of high workload and limited resources — they may be more likely to be effective. Different strategies may produce effects via different mechanisms. Several reviews have reported that multi-component interventions are the most effective in changing test-ordering behaviour.<sup>14,16–18</sup> Complex interventions may act to exert greater pressure to change perceptions of laboratory testing, provoking multiple mechanisms that 'work' for different individuals and reflecting the wide variation in factors that influence clinicians' testing practice.

Where interventions are unsuccessful, the authors' programme theory suggests that this may reflect a failure to convince clinicians that laboratory testing is important, and/or that the time and resources for thorough decision making were not available.<sup>20,21</sup> It may, therefore, be instructive for future evaluations of interventions to attempt to uncover which mechanisms and associated intervention strategies have been effective in which contexts, and to include an assessment of outcomes relating to clinicians' perceptions of the relative triviality or importance of laboratory testing and workloads.

Finally, the complexity of test-ordering practice requires that future research should consider the potential unintended consequences of interventions designed to change test-ordering practice. For



example, where intervention design aims to reduce test ordering for social or strategic purposes, evaluations should ensure that potential trade-offs in efficiency and thoroughness are considered; for example, what are the side effects (in relation to clinicians' workloads, as well as for patients) of preventing or encouraging clinicians to avoid deploying tests in this way? Approaches that permit theory-informed intervention design and evaluations that take account of complexity, differences in context and unintended consequences are recommended, especially realist evaluation.<sup>68</sup>

### Funding

CD is funded by a National Institute of Health Research (NIHR) Research Methods Programme Systematic Review Fellowship (NIHR-RM-SR-2017-08-018). GW's salary is partly supported by the Evidence Synthesis Working Group of the NIHR School of Primary Care (Project Number 390). This (publication/paper/report) presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

### Ethical approval

Not applicable; review of secondary data.

### Provenance

Freely submitted; externally peer reviewed.

### Acknowledgements

The authors offer their sincere thanks to their stakeholder group for their time and invaluable input at several stages during this review.

### Competing interests

CD and GW are both members of the Royal College of General Practitioners (UK) Overdiagnosis and Overtreatment Group. GW is an NHS general practitioner and deputy chair of the NIHR Health Technology Assessment Prioritisation Committee: Integrated Community Health and Social Care Panel (A).

## References

1. Public Health England, NHS RightCare. The NHS atlas of variation in diagnostic services. Reducing unwarranted variation to increase value and improve quality. 2013; <https://fingertips.phe.org.uk/profile/atlas-of-variation> (accessed 11 Nov 2020).
2. Public Health England, NHS RightCare. The 2nd atlas of variation in NHS diagnostic services in England. Reducing unwarranted variation to improve health outcomes and value. 2017; <http://fingertips.phe.org.uk/profile/atlas-of-variation> (accessed 11 Nov 2020).
3. O'Sullivan JW, Albasri A, Nicholson BD, et al. Overtesting and undertesting in primary care: a systematic review and meta-analysis. *BMJ Open* 2018; **8**(2): e018557. DOI: <https://doi.org/10.1136/bmjopen-2017-018557>
4. Zhi M, Ding EL, Theisen-Toupal J, et al. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. *PLoS One* 2013; **8**(11): e78962. DOI: <https://doi.org/10.1371/journal.pone.0078962>
5. Busby J, Schroeder K, Woltersdorf W, et al. Temporal growth and geographic variation in the use of laboratory tests by NHS general practices: using routine data to identify research priorities. *Br J Gen Pract* 2013; **63**(609): e256–e266. DOI: <https://doi.org/10.3399/bjgp13X665224>
6. O'Sullivan JW, Stevens S, Hobbs FDR, et al. Temporal trends in use of tests in UK primary care, 2000–15: retrospective analysis of 250 million tests. *BMJ* 2018; **363**: k4666. DOI: <https://doi.org/10.1136/bmj.k4666>
7. NHS Improvement. Pathology networks. 2017; <https://improvement.nhs.uk/resources/pathology-networks/> (accessed 11 Nov 2020).
8. Lord Carter of Coles. Report of the review of NHS pathology services in England. 2006; <https://www.networks.nhs.uk/nhs-networks/peninsula-pathology-network/documents/CarterReviewPathologyReport.pdf> (accessed 11 Nov 2020).
9. NHS England. National pathology programme. Digital first: clinical transformation through pathology innovation. 2014; <https://www.england.nhs.uk/wp-content/uploads/2014/02/pathol-dig-first.pdf> (accessed 11 Nov 2020).
10. Hofmann B, Welch HG. New diagnostic tests: more harm than good. *BMJ* 2017; **358**: j3314. DOI: <https://doi.org/10.1136/bmj.j3314>
11. McCaffery KJ, Jansen J, Scherer LD, et al. Walking the tightrope: communicating overdiagnosis in modern healthcare. *BMJ* 2016; **352**: i348. DOI: <https://doi.org/10.1136/bmj.i348>
12. Fisher ES, Welch HG. Avoiding the unintended consequences of growth in medical care: how might more be worse? *JAMA* 1999; **281**(5): 446–453. DOI: <https://doi.org/10.1001/jama.281.5.446>

13. Harris RP, Sheridan SL, Lewis CL, et al. The harms of screening: a proposed taxonomy and application to lung cancer screening. *JAMA Intern Med* 2014; **174**(2): 281–285. DOI: <https://doi.org/10.1001/jamainternmed.2013.12745>
14. Solomon DH, Hashimoto H, Daltroy L, et al. Techniques to improve physicians' use of diagnostic tests. *JAMA* 1998; **280**(23): 2020. DOI: <https://doi.org/10.1001/jama.280.23.2020>
15. Main C, Moxham T, Wyatt JC, et al. Computerised decision support systems in order communication for diagnostic, screening or monitoring test ordering: systematic reviews of the effects and cost-effectiveness of systems. *Health Technol Assess* 2010; **14**(48): 1–227. DOI: <https://doi.org/10.3310/hta14480>
16. Smellie WSA. Demand management and test Request rationalization. *Ann Clin Biochem* 2012; **49**(Pt 4): 323–336. DOI: <https://doi.org/10.1258/acb.2011.011149>
17. Cadogan SL, Browne JP, Bradley CP, et al. The effectiveness of interventions to improve laboratory requesting patterns among primary care physicians: a systematic review. *Implement Sci* 2015; **10**: 167. DOI: <https://doi.org/10.1186/s13012-015-0356-4>
18. Kobewka DM, Ronksley PE, McKay JA, et al. Influence of educational, audit and feedback, system based, and incentive and penalty interventions to reduce laboratory test utilization: a systematic review. *Clin Chem Lab Med* 2015; **53**(2): 157–183. DOI: <https://doi.org/10.1515/cclm-2014-0778>
19. Thomas RE, Vaska M, Naugler C, et al. Interventions at the laboratory level to reduce laboratory test ordering by family physicians: systematic review. *Clin Biochem* 2015; **48**(18): 1358–1365. DOI: <https://doi.org/10.1016/j.clinbiochem.2015.09.014>
20. Thomas RE, Vaska M, Naugler C, et al. Interventions to educate family physicians to change test ordering: systematic review of randomized controlled trials. *Acad Pathol* 2016; **3**: 237428951663347. DOI: <https://doi.org/10.1177/2374289516633476>
21. Zhelev Z, Abbott R, Rogers M, et al. Effectiveness of interventions to reduce ordering of thyroid function tests: a systematic review. *BMJ Open* 2016; **6**(6): e010065. DOI: <https://doi.org/10.1136/bmjopen-2015-010065>
22. Delvaux N, Van Thienen K, Heselmans A, et al. The effects of computerized clinical decision support systems on laboratory test ordering: a systematic review. *Arch Pathol Lab Med* 2017; **141**(4): 585–595. DOI: <https://doi.org/10.5858/arpa.2016-0115-RA>
23. Maillet Éric, Paré G, Currie LM, et al. Laboratory testing in primary care: a systematic review of health it impacts. *Int J Med Inform* 2018; **116**: 52–69. DOI: <https://doi.org/10.1016/j.ijmedinf.2018.05.009>
24. Pawson R. Realist methodology: the building blocks of evidence. In: *Evidence-based Policy*. London: Sage Publications; 2006. 17–37.
25. Duddy C, Wong G. Explaining variations in test ordering in primary care: protocol for a realist review. *BMJ Open* 2018; **8**(9): e023117. DOI: <https://doi.org/10.1136/bmjopen-2018-023117>
26. Pawson R. Realist synthesis: new protocols for systematic review. In: *Evidence-based Policy*. London: Sage Publications; 2006. 73–104.
27. The RAMESES Project. Quality standards for realist synthesis. 2014; [https://www.ramesesproject.org/Standards\\_and\\_Training\\_materials.php](https://www.ramesesproject.org/Standards_and_Training_materials.php) (accessed 21 Jan 2021).
28. Wong G, Greenhalgh T, Westhorp G, et al. RAMESES publication standards: realist syntheses. *BMC Med* 2013; **11**(1): 21. DOI: <https://doi.org/10.1186/1741-7015-11-21>
29. Tracy CS, Dantas GC, Moineddin R, et al. Contextual factors in clinical decision making: national survey of Canadian family physicians. *Can Fam Physician* 2005; **51**: 1106–1107.
30. Litchfield IJ, Lilford RJ, Bentham LM, et al. A qualitative exploration of the motives behind the decision to order a liver function test in primary care. *Qual Prim Care* 2014; **22**(4): 201–210.
31. Opdal Per Øystein, Meland E, Hjörleifsson S. Dilemmas of medical overuse in general practice — A focus group study. *Scand J Prim Health Care* 2019; **37**(1): 135–140. DOI: <https://doi.org/10.1080/02813432.2019.1569370>
32. van der Weijden T, van Bokhoven MA, Dinant G-J, et al. Understanding laboratory testing in diagnostic uncertainty: a qualitative study in general practice. *Br J Gen Pract* 2002; **52**(485): 974–980.
33. Watson J, de Salis I, Banks J, et al. What do tests do for doctors? A qualitative study of blood testing in UK primary care. *Fam Pract* 2017; **34**(6): 735–739. DOI: <https://doi.org/10.1093/fampra/cmz051>
34. Sabbatini AK, Tilburt JC, Campbell EG, et al. Controlling health costs: physician responses to patient expectations for medical care. *J Gen Intern Med* 2014; **29**(9): 1234–1241. DOI: <https://doi.org/10.1007/s11606-014-2898-6>
35. van Bokhoven MA, Koch H, van der Weijden T, et al. The effect of watchful waiting compared to immediate test ordering instructions on general practitioners' blood test ordering behaviour for patients with unexplained complaints; a randomized clinical trial (ISRCTN55755886). *Implement Sci* 2012; **7**: 29. DOI: <https://doi.org/10.1186/1748-5908-7-29>
36. Powell AA, Bloomfield HE, Burgess DJ, et al. A conceptual framework for understanding and reducing overuse by primary care providers. *Med Care Res Rev* 2013; **70**(5): 451–472. DOI: <https://doi.org/10.1177/1077558713496166>
37. Deyo RA. Cascade effects of medical technology. *Annu Rev Public Health* 2002; **23**: 23–44. DOI: <https://doi.org/10.1146/annurev.publhealth.23.092101.134534>
38. Hollnagel E. *The ETTO Principle: Efficiency-Thoroughness Trade-Off: Why Things That Go Right Sometimes Go Wrong*. Farnham: Ashgate Publishing; 2009.
39. Pathirana T, Clark J, Moynihan R. Mapping the drivers of overdiagnosis to potential solutions. *BMJ* 2017; **358**: j3879. DOI: <https://doi.org/10.1136/bmj.j3879>
40. Hobbs FDR, Bankhead C, Mukhtar T, et al. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14. *Lancet* 2016; **387**(10035): 2323–2330. DOI: [https://doi.org/10.1016/S0140-6736\(16\)00620-6](https://doi.org/10.1016/S0140-6736(16)00620-6)

41. Croxson CH, Ashdown HF, Hobbs FR. GPs' perceptions of workload in England: a qualitative interview study. *Br J Gen Pract* 2017; **67(655)**: e138–e147. DOI: <https://doi.org/10.3399/bjgp17X688849>
42. Fisher RF, Croxson CH, Ashdown HF, et al. GP views on strategies to cope with increasing workload: a qualitative interview study. *Br J Gen Pract* 2017; **67(655)**: e148–e156. DOI: <https://doi.org/10.3399/bjgp17X688861>
43. Grant S, Guthrie B. Efficiency and thoroughness trade-offs in high-volume organisational routines: an ethnographic study of prescribing safety in primary care. *BMJ Qual Saf* 2018; **27(3)**: 199–206. DOI: <https://doi.org/10.1136/bmjqs-2017-006917>
44. Duncan P, Cabral C, McCahon D, et al. Efficiency versus thoroughness in medication review: a qualitative interview study in UK primary care. *Br J Gen Pract* 2019; **69(680)**: e190–e198. DOI: <https://doi.org/10.3399/bjgp19X701321>
45. Sood R, Sood A, Ghosh AK. Non-evidence-based variables affecting physicians' test-ordering tendencies: a systematic review. *Neth J Med* 2007; **65(5)**: 167–177.
46. Whiting P, Toerien M, de Salis I, et al. A review identifies and classifies reasons for ordering diagnostic tests. *J Clin Epidemiol* 2007; **60(10)**: 981–989. DOI: <https://doi.org/10.1016/j.jclinepi.2007.01.012>
47. Grant S, Checkland K, Bowie P, et al. The role of informal dimensions of safety in high-volume organisational routines: an ethnographic study of test results handling in UK general practice. *Implement Sci* 2017; **12(1)**: 56. DOI: <https://doi.org/10.1186/s13012-017-0586-8>
48. Bate L, Hutchinson A, Underhill J, et al. How clinical decisions are made. *Br J Clin Pharmacol* 2012; **74(4)**: 614–620. DOI: <https://doi.org/10.1111/j.1365-2125.2012.04366.x>
49. Balla JL, Heneghan C, Glasziou P, et al. A model for reflection for good clinical practice. *J Eval Clin Pract* 2009; **15(6)**: 964–969. DOI: <https://doi.org/10.1111/j.1365-2753.2009.01243.x>
50. Gabbay J, le May A, Al M. Evidence based guidelines or collectively constructed "mindlines?" Ethnographic study of knowledge management in primary care. *BMJ* 2004; **329(7473)**: 1013. DOI: <https://doi.org/10.1136/bmj.329.7473.1013>
51. Shalev V, Chodick G, Heymann AD. Format change of a laboratory test order form affects physician behavior. *Int J Med Inform* 2009; **78(10)**: 639–644. DOI: <https://doi.org/10.1016/j.ijmedinf.2009.04.011>
52. O'Connor PJ, Crain AL, Rush WA, et al. Impact of an electronic medical record on diabetes quality of care. *Ann Fam Med* 2005; **3(4)**: 300–306. DOI: <https://doi.org/10.1370/afm.327>
53. Steele AW, Eisert S, Witter J, et al. The effect of automated alerts on provider ordering behavior in an outpatient setting. *PLoS Med* 2005; **2(9)**: e255. DOI: <https://doi.org/10.1371/journal.pmed.0020255>
54. Sequist TD, Zaslavsky AM, Marshall R, et al. Patient and physician reminders to promote colorectal cancer screening: a randomized controlled trial. *Arch Intern Med* 2009; **169(4)**: 364–371. DOI: <https://doi.org/10.1001/archinternmed.2008.564>
55. MacLean CD, MacCaskey M, Littenberg B. Improving testing for proteinuria in diabetes using decision support: role of laboratory ordering systems. *Lab Med* 2013; **44(4)**: 353–357. DOI: <https://doi.org/10.1309/LMCFWN1FX1QZWRQ2>
56. Kenealy T, Arroll B, Petrie KJ. Patients and computers as reminders to screen for diabetes in family practice. Randomized-controlled trial. *J Gen Intern Med* 2005; **20(10)**: 916–921. DOI: <https://doi.org/10.1111/j.1525-1497.2005.0197.x>
57. van Wyk JT, van Wijk MAM, Sturkenboom MCJM, et al. Electronic alerts versus on-demand decision support to improve dyslipidemia treatment: a cluster randomized controlled trial. *Circulation* 2008; **117(3)**: 371–378. DOI: <https://doi.org/10.1161/CIRCULATIONAHA.107.697201>
58. van Wijk MAM, van der Lei J, Mosseveld M, et al. Compliance of general practitioners with a guideline-based decision support system for ordering blood tests. *Clin Chem* 2002; **48(1)**: 55–60. DOI: <https://doi.org/10.1093/clinchem/48.1.55>
59. Thomas RE, Croal BL, Ramsay C, et al. Reducing inappropriate laboratory testing through provider education and feedback. *J Clin Outcomes Manag* 2006; **13(9)**: 472–477.
60. Lo HG, Matheny ME, Seger DL, et al. Impact of non-interruptive medication laboratory monitoring alerts in ambulatory care. *J Am Med Inform Assoc* 2009; **16(1)**: 66–71. DOI: <https://doi.org/10.1197/jamia.M2687>
61. Palen TE, Raebel M, Lyons E, et al. Evaluation of laboratory monitoring alerts within a computerized physician order entry system for medication orders. *Am J Manag Care* 2006; **12(7)**: 389–395.
62. Matheny ME, Sequist TD, Seger AC, et al. A randomized trial of electronic clinical reminders to improve medication laboratory monitoring. *J Am Med Inform Assoc* 2008; **15(4)**: 424–429. DOI: <https://doi.org/10.1197/jamia.M2602>
63. Sequist TD, Gandhi TK, Karson AS, et al. A randomized trial of electronic clinical reminders to improve quality of care for diabetes and coronary artery disease. *J Am Med Inform Assoc* 2005; **12(4)**: 431–437. DOI: <https://doi.org/10.1197/jamia.M1788>
64. Marewski JN, Gigerenzer G. Heuristic decision making in medicine. *Dialogues Clin Neurosci* 2012; **14(1)**: 77–89.
65. Croskerry P. Achieving quality in clinical decision making: cognitive strategies and detection of bias. *Acad Emerg Med* 2002; **9(11)**: 1184–1204. DOI: <https://doi.org/10.1197/aemj.9.11.1184>
66. Elwenspoek M, Patel R, Watson J, et al. What is the evidence behind guideline recommendations to monitor chronic diseases in UK primary care? *Br J Gen Pract* 2019; **69(suppl 1)**: bjgp19X702917. DOI: <https://doi.org/10.3399/bjgp19X702917>
67. Lang T, Croal B. *National minimum retesting intervals in pathology*. London: The Royal College of Pathologists; 2015.
68. Wong G, Westhorp G, Manzano A, et al. RAMESES II reporting standards for realist evaluations. *BMC Med* 2016; **14(1)**: 96. DOI: <https://doi.org/10.1186/s12916-016-0643-1>

69. Booth A, Harris J, Croot E, et al. Towards a methodology for cluster searching to provide conceptual and contextual "richness" for systematic reviews of complex interventions: case study (CLUSTER). *BMC Med Res Methodol* 2013; **13**(1): 118. DOI: <https://doi.org/10.1186/1471-2288-13-118>
70. Booth A. Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. *Syst Rev* 2016; **5**(1): 74. DOI: <https://doi.org/10.1186/s13643-016-0249-x>
71. Booth A, Wright J, Briscoe S. Scoping and searching to support realist approaches. In: Emmel N, Greenhalgh J, Manzano A, et al, eds. *Doing Realist Research*. London: Sage Publications; 2018. 148–165.
72. Wong G. Data gathering in realist reviews: looking for needles in haystacks. In: Emmel N, Greenhalgh J, Manzano A et al, eds. *Doing Realist Research*. London: Sage Publications; 2018. 131–146.