International Journal of Technology Assessment in Health Care, 28:3 (2012), 211–219.

doi:10.1017/S0266462312000402

ENGAGING CLINICIANS IN EVIDENCE-BASED DISINVESTMENT: ROLE AND PERCEPTIONS OF EVIDENCE

Amber M. Watt, Cameron D. Willis, Katherine Hodgetts
The University of Adelaide

Janet E. Hiller
Australian Catholic University

Adam G. Elshaug
The University of Adelaide; Harvard Medical School
email: elshaug@hcp.med.harvard.edu

For The ASTUTE Health Study Group

Objectives: The aim of this study was to determine how evidence from systematic review (SR) is perceived and negotiated by expert stakeholders in considering a technology for potential disinvestment.
Methods: An evidence-informed stakeholder engagement examined results from a diagnostic accuracy SR of vitamin B_{12} and folate tests. Pathologists deliberated around the SR findings to generate an informed contribution to future policy for the funding of B_{12} and folate tests. Deliberations were transcribed and subject to qualitative analysis.
Results: Pathologists did not engage with findings from the SR in depth; rather they sought to contest the terms of the problem driving the review and attempted to reframe it. Pathologists questioned the usefulness of SR outcomes given the variable definitions of B_{12} deficiency and deferred addressing disinvestment options specifically pertaining to B_{12} testing. However, *folate* testing was proffered as a potential disinvestment candidate, based upon pathologists' definition of "appropriate" evidence beyond the bounds of the SR.
Conclusions: The value of SR to informing disinvestment deliberations by expert stakeholders may be a function of timing as well as content. Engagement of stakeholders in co-produced evidence may be required at two levels: (i) Early in the synthesis phase to help shape the SR and harmonize expert views with the available evidence (including gaps); (ii) Collaboration in primary research to fill evidence-gaps thus supporting evidence-based disinvestment. Without this, information asymmetry between clinically engaged experts and decision makers may preclude the collaborative, informed, and technical discussions required to generate productive policy change.

Keywords: Disinvestment, Qualitative research, Systematic review, Evidence-based practice, Health policy

Health policy making in Australia, as in other Western democratic contexts, has increasingly been characterized as a process of evidence appraisal (12). Mirroring the developing prominence of evidence-based medicine, "evidence-informed health policy" has emerged as a dominant evaluative framework in which decisions routinely foreground the outcomes of primary research, systematic reviews (SR), and health technology assessments (HTA).

While some practitioners critique the notion that "evidence-based healthcare. . .can make an unproblematic transformation into evidence-based policy" (11), a growing acknowledgement that decision-making processes should be both explicit and public has fuelled the evidence-based policy movement. Such concerns are particularly salient with regard to healthcare "disinvestment" – policy reform oriented to supporting safe and effective care within sustainable health systems (7).

We thank the participants in the stakeholder engagements; without their generous contributions this research would not have been possible. Prof Hiller and Dr Elshaug are co-chief-investigators of The ASTUTE Health study which is funded by the National Health and Medical Research Council (NHMRC) of Australia (project ID 565327). Additional members of The ASTUTE Health study group are: Prof Annette Braunack-Mayer, Dr John R. Moss, Dr Heather Buchan and Dr Janet Wale. Drs Elshaug and Willis hold NHMRC Fellowships (ID 627061 and 1013165, respectively).

In Australia, SR evidence forms the cornerstone of HTA and related policy processes. Within a disinvestment framework, understanding how SR findings fit within the rubric of "evidence," how they are used to inform decisions, and how this evidence is negotiated by expert clinical stakeholders, is integral to the establishment, evaluation, and success of reform. Using a case study, this investigation describes an example of evidence-informed debate around potential disinvestment policy, situated within a case study of pathology testing for vitamin B_{12} and folate. More specifically, we address the role of SR in informing such debates, and the negotiation of SR evidence by an expert group of clinical stakeholders.

Systematic Review in a Health Policy Context

Systematic reviews have become central starting points to framing evidence-informed health *investment* decisions made at local, provincial/territorial, and federal levels in several international jurisdictions (for examples, see 1;16;19;21).

However, determining the role of SR in the context of *disinvestment* decisions has been less straightforward. Early work in the Australian context has suggested broad stakeholder support for disinvestment processes to be aligned with existing evidence-based policy processes for investment decisions (6), while Spanish work in this area has also emphasized the use

of SR in disinvestment decisions (9). The United Kingdom has focused more broadly on using guidelines as the basis for disinvestment recommendations (including those based on Cochrane Collaboration SRs), while Program Budgeting and Marginal Analysis (PBMA) approaches, among others, have been applied in Canadian provinces and regions (9).

For *new* and *emerging* technologies and services, comprehensive SRs are an integral part of HTA but they alone do not constitute an HTA. For disinvestment health policy decision making the precise role of SR requires investigation, given certain conditions unique to disinvestment policy, including: contention around the identification of candidate technologies; the apparent greater "strength" of evidence required for a disinvestment decision; the contestability of evidence; and the role of stakeholders in these decisions.

To inform further the pragmatic application of SR methods to disinvestment decisions, this study addresses broad questions around the role of SR within disinvestment policy processes. The discussion focuses on how clinical stakeholders perceived and negotiated SR evidence within a specific case study with potential for disinvestment: pathology testing for vitamin B_{12} and folate.

METHODS

Selection of Case Study

Lead members of this research group have proposed criteria by which to "flag" potential technologies that might warrant further investigation under disinvestment programs (8). These criteria include among others: geographic variations in use; temporal variations in volume; public interest or controversy; or nomination from clinical groups. This process was adapted for topic selection in the current study, resulting in a wide range of candidates being flagged. Tests for investigating vitamin B_{12} and folate levels—most frequently ordered conjointly in Australia—were nominated by clinical experts to this research group. Further scoping revealed that they met several additional "flag" criteria from this group's identification and prioritization frameworks (8); a full account of this is recorded in a separate manuscript (Willis et al., unpublished data 2012). Briefly, since 2001, usage of a combined vitamin B_{12} and folate test through Australia's universal health insurance scheme, Medicare, has experienced an annual average growth of 21.78 percent (15) in contrast to 6.3 percent for all pathology services (14). International differences in recommended test usage are also apparent (3;10;17;20), while usage in some clinical scenarios appears to conflict with available clinical guidelines (2). Finally, given this was a primary research study, it was decided by the lead investigators that these tests offered characteristics (e.g., low level of controversy, high volume, low cost per-item) which would balance the other case study chosen for investigation, assisted reproductive technologies.

Systematic Review

Diagnostic accuracy (a central marker of the effectiveness of diagnostic testing) formed the foundations of a SR on the effectiveness of vitamin B_{12} and folate tests in determining the presence or absence of deficiencies (22). A protocol for guiding this SR was developed in consultation with a clinical pathologist, health services researchers, epidemiologists, health economists and experts in HTA. The SR aimed to synthesize published evidence relating to the diagnostic accuracy of the three most commonly used tests for investigating vitamin B_{12} and folate deficiency: serum B_{12}, serum folate, and red cell folate assays.

Results from the SR and meta-analysis suggested that the serum B_{12} assay has limited ability to identify patients with symptoms amenable to vitamin supplementation, noting poor positive (PLR) and negative (NLR) likelihood ratios across patient sub-groups. Similar results were found for serum folate assays, while there was insufficient evidence available to complete a similar analysis for red cell folate assays. As noted in the review, the existing evidence base is hampered by lack of an internationally recognized "gold standard" for diagnosing either vitamin B_{12} or folate deficiencies; imperfect reference standard tests; and inconsistent "cut-off" values used to distinguish deficient from nondeficient patients (22).

Stakeholder Engagement

Following completion of the SR, a series of iterative stakeholder engagements was undertaken, involving test requesters (primary care physicians and specialists), test providers (pathologists) and key decision makers (federal government policy advisors). Details of this engagement process are provided elsewhere (Watt et al., unpublished data, 2012). Briefly, each group was consulted separately in professionally facilitated sessions that sought to explore reasons for the changing usage of B_{12}/folate tests, to discuss the evidence arising from the SR, and to determine potential avenues for reform (see Figure 1).

Nominations for participants within the test provider group (the focus of this manuscript) were sought from key stakeholder bodies in the field of pathology (see Table 1). Nine nominees and key opinion leaders (pathologists) took part in the engagement held in Sydney, New South Wales in May 2011; three females and six males. During this 4-hour session, participants were presented with the outcomes of the SR, supplemented with information modules pertaining to relevant health funding policy and economic considerations to support decision making (see Figure 1). Following deliberation around this evidence, participants were invited to discuss their experiential knowledge and perspectives on the issue, before formulating a response to the question "What should be considered when making decisions about how much we should publicly subsidies B_{12}/folate pathology tests?". After obtaining consent from participants, deliberations were audio-recorded and transcribed verbatim by an experienced stenographer.

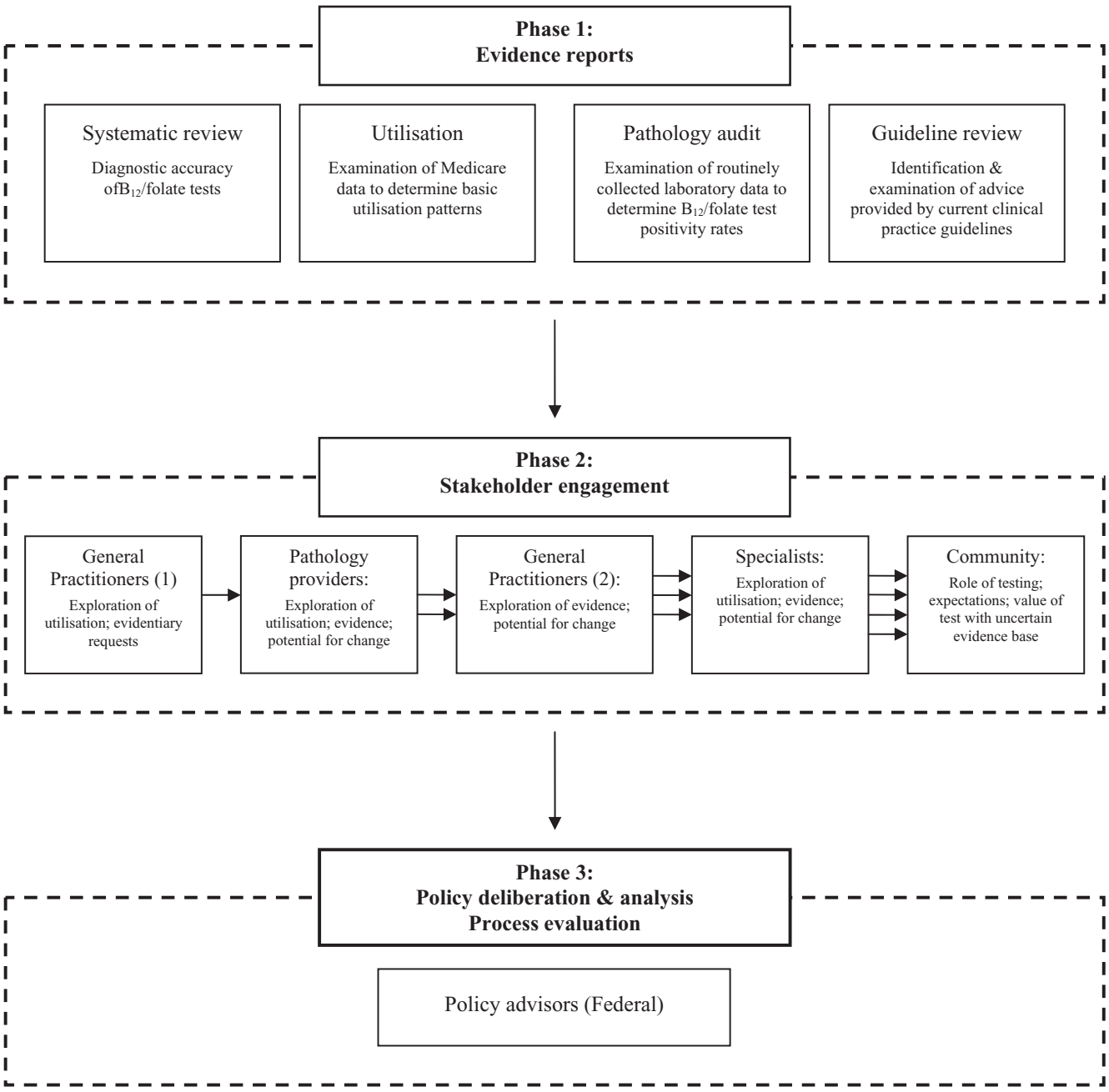

Figure 1. Evidence generation and stakeholder engagement process for vitamin B_{12}/folate pathology testing case study.

Data Analysis

The theoretical perspective of this study is broadly social constructionist, our analysis grounded in the assumption that talk reflects both the interactional context in which it is generated, as well as broader patterns of sense making that can have both ideological and material consequences (5).

Our focus is on how expert stakeholders (pathologists) made sense of "evidence" within a data-informed discussion on a disinvestment topic. In particular, we attend to participants' framings of "what counts" as appropriate evidence for disinvestment, the language used to justify their claims, and the implications of these arguments for engaged policy processes around disinvestment more broadly.

Transcripts of the stakeholder engagement followed the model for thematic analysis as outlined by Braun and Clarke (4). Through a process of open coding, the whole data corpus was worked through to build up a file of instances referencing "evidence" with regard to the case study of B_{12}/folate testing. The main repetitive patterns evident across these accounts were identified and coded (axial coding). The whole corpus was then re-read in light of these codes, and over-arching themes (constructed so as to emphasize internal convergence and external

Table 1. Composition and Characteristics of Participants (Pathologists)

Target source(s)	Inclusion criteria	Characteristics
Nominations from national public & private pathology interest groups[a]; purposive sampling of key opinion leaders; snowballing	Participants must be working full-time or part-time in administrative, management or laboratory roles in private or public pathology providers in Australia	6 males; 3 females *Role & practice:* Public pathology: 4/9 Private pathology: 5/9 *Geographic regions:* New South Wales: 4/9 Victoria: 2/9 Queensland: 1/9 Tasmania: 1/9National: 1/9

[a] Including National Coalition of Public Pathology, Australian Association of Pathology Practices; Australasian Association of Clinical Biochemists; Healthscope; Sonic; and Primary Health.

divergence) were subsequently identified (4). These themes, and representative data extracts, are outlined (in no particular order) in Table 2, and we analyze each in turn to elaborate their implications with regard to questions of evidence in disinvestment policy.

In representing extracts of policy advisors' talk, the transcription protocol is as follows: [] indicates clarification of meaning; . . . indicates words omitted for ease of reading; (. . .) indicates inaudible speech.

RESULTS

Pathologists questioned the usefulness of the outcomes of the SR presented to them at the outset of the engagement, and the appropriateness of using systematic reviewing *per se* as a means of informing funding policy around vitamin B_{12}/folate testing. More specifically, pathologists: (1) *Contested the terms of the problem* driving the review, arguing that the B_{12} element of the combined B_{12}/folate test (but not the folate) does not represent a legitimate candidate for disinvestment discussion. Note: that pathologists did nominate this as a case study points to a lack of agreement within the group (see point 2). (2) *Reframed the terms of the problem* that ought to have driven the review, suggesting that only folate testing is a suitable disinvestment focus (given that mandatory fortification has made the identification of folate deficiency rare), and that a clear definition of deficiency must be established before the effectiveness of vitamin B_{12} testing may be usefully investigated using SR methods.

We address each of these arguments, and their consequences, in turn.

Contesting the Terms of the Problem

A central theme in the pathologists' deliberations was that the usage of testing for levels of vitamin B_{12} (as opposed to folate) is currently *appropriate*, a claim that drew into question the fundamental value of undertaking a SR in this area.

Table 2 presents a summary of participants' arguments in support of this position. These arguments have significant implications, not only for this case study, but for the more general role of SR within disinvestment-oriented policy processes. For example, arguments 1–3 undermined the driving rationale of a "review for disinvestment" in this area. Holding that B_{12} testing remains imperative for a range of presentations, and that utilisation reflects genuine "medical issues" rather than the vested "interest" of laboratories, pathologists rejected the notion that this item should be subject to disinvestment discussion *per se*. Although a consensually held position across the engagement, it must be noted that this argument is at odds with the original nomination of these B_{12}/folate tests by other members of the pathology community as warranting disinvestment consideration.

Within a disinvestment process, participants' rejection of a technology as a legitimate disinvestment candidate represents a significant barrier to productive stakeholder engagement. In this case, a rejection of what was perceived to be the guiding problem ("B_{12} test utilisation is inappropriate, thus we should consider disinvestment") forestalled potential conversations with pathologists around the effectiveness of this test as indicated by the SR. The notion that disinvestment might represent a more nuanced outcome than blunt rationing (to fund or not to fund) was not entertained within the arguments outlined. For example, pathologists' claim that B_{12} testing remains imperative for particular populations—a notion that could have supported calls for "restrictions in eligibility" for this test to those specific groups—served in the examples cited in Table 2 to argue in favor of simply leaving the test alone. Stakeholder interpretations of review outcomes may be colored by "what stands to be lost", and a defensive stance around accountability for perceived overuse (see argument 3) may preclude discussion around more fine-grained disinvestment options.

Arguments 4–5 presented in Table 2 also have implications for the use of SRs in disinvestment contexts. With regard to the case study of B_{12}, these claims suggested that as long as deficiencies continue to be found, and there exists no better available assay, we have no choice but to "use the test we have." These arguments served to position the outcomes of the SR as "purely academic," and that indications of a test's accuracy are useful only when alternative tests are available. In turn, these claims precluded engagement with disinvestment concerns around quality as indicated by the review (e.g., at what point might it be wise to restrict or remove a test even in the absence of alternatives?). At the same time, the superiority of an alternative assay (holotranscobalamin) was asserted within these arguments without empirical evidence of its superior accuracy. Each of these claims suggests the importance of explicitly discussing issues of evidence (what counts and how to count it) in both the framing and interpretation of SRs for disinvestment.

Argument 6, holding that any partial disinvestment/restriction to pathology testing will represent additional costs,

Table 2. Quotes Illustrating Pathologists' Position That B_{12} Testing Is Appropriate

	B12 test utilisation is appropriate. . .	Illustrative quotes
1	. . .For specific populations and presentations	"A number of [reasons for testing are] quite appropriate, like dementia, tiredness, and vegetarians, and alcohol use" (Participant 4) "testing [is] important in coeliac patients and other patients with malabsorption, so that's the high risk patient" (Participant 8)
2	. . .Because there are legitimate reasons utilisation has grown	"There is much greater awareness in the community leading to B_{12} testing. There's been some very interesting papers recently, which they're obviously reading . . . that have shown that if dementia and B_{12} are linked and it's treated within 2 years, it's fully reversible . . . and if it's not treated within 2 years it's permanent" (Participant 6) "People are becoming a little bit more aware of screening for causes of cognitive decline" (Participant 9) "We've got this increase in [the prescription of] Metformin for Type 2 diabetes and that has a significant impact on the way B_{12} works" (Participant 9) "I studied lacto-ovo vegetarians from 1972 to 2002 and there's a direct correlation between a reduction in egg and milk consumption [and a] reduction in serum B_{12} levels, and this was probably due to medical education relating to cholesterol" (Participant 7) " there is [no longer] any ancillary testing and then you have to repeat your primary test to see if it is a different result" (Participant 1)
3	. . .Because it is not driven by local interest	"The growth [in B_{12} test utilisation] has been very similar [across] the states. . . .If you look at PSA [prostate specific antigen testing] there are tenfold variations from one state to another in per capita usage. That tells me [B_{12} testing is] not being driven by local interest, it's being driven by a general medical issue . . . it's not something the laboratories are contributing to"
4	. . . Because we continue to find deficiencies	"The relative incidence [of deficiency] hasn't changed although the absolute numbers being identified, and disease, obviously increases with the numbers that are requested" (Participant 1)
5	. . .Because we do not have better options	I think we are using the wrong test . . . holotranscobalamin is actually a much better assay because it reflects the fraction that goes directly inside the cells. (Participant 4) "Just the mechanics with the holotranscobalamin assay, I mean most laboratories say we don't do it because we can't afford it or we don't have the machines to do it . . . It's going to be one to two years before it's available to all laboratories so it's a bit hard, we use what we can even though we'd like everyone to use active B_{12} in practice" (Participant 8)
6a	. . .Because it costs more (to the patient) to restrict testing	"It seems like it [B_{12} testing] is money well spent when you compare it to the cost of inappropriately labelling someone as being B_{12} deficient and a lifetime of injections they didn't need" (Participant 9) "I think the number, the dollars per low B_{12} may not be overspending public money. I think it can be very significant to identify a single individual who's B_{12} deficient . . . in terms of impact on your life and wellbeing" (Participant 6)
6b	. . .Because it costs more (to the system) to restrict testing	"[When test restrictions are enacted, labs] spend heaps and heaps of money trying to work out which one they can bill or not, and it might be more efficient for us just to do them all" (Participant 8) "It might be more costly to do more. We want . . . to point out to the government all these costs of regulation so every time . . . anybody imposes anything, there's an additional cost" (Participant 2)

suggests an inevitable resistance to SR outcomes in this context. Without unequivocal evidence of harm, inaccuracy or superior alternatives it seems that review outcomes indicating support for "refinements to indications" may always be susceptible to arguments foregrounding the "burden of regulation." This finding indicates the importance of enhancing SR outcomes with economic analyses at an early point in an engaged policy process in order that discussion of the costs and benefits of disinvestment options may remain appropriately evidence-based.

Evidence and Judgment

As indicated above, there was general agreement among pathologists that the holotranscobalamin assay (as yet not publicly funded in Australia, nor widely available) represents a superior B_{12} test than that currently used in most Australian laboratories. Interestingly, the supposed superior accuracy of holotranscobalamin was not taken as evidence that disinvestment from the current B_{12} item may be justified. On the contrary; participants argued that inaccuracy is inherent in *all* testing, an issue

suggesting that decisions about what constitutes "appropriate test utilization" rely, not on evidence such as that generated from SR, but on "judgments" around *sensitivity versus specificity*, the *value of screening*, and the *definition of reference standards and ranges* for deficiency; all questions of values.

Sensitivity Versus Specificity

Participant (P) 6: *Even that test [holotranscobalamin], while it's a good test, the predictive value will be limited. You'll still have false positives and false negatives no matter what you do.*

The Value of Screening

P9: *The difference between therapeutics and diagnostics is absolutely enormous . . . [In therapeutics] you can measure side effects, efficacy and costs much more easily than in diagnostics, where essentially we're offering screening for disease and if you're doing your job properly most of the results are going to be normal. It's a philosophical question, how much investment the community wants to make in screening.*

The Definition of "Reference Standards" and "Reference Ranges"

P6: *The whole study you [the research team] have done is predicated on reference ranges and I know some people who have used MCV [mean corpuscular volume] as the standard. There's no problem with B_{12} until the MCV is elevated. Well, you miss 80 percent of all cases if that's the reference standard you're using.*

P2: *In the case of B_{12}, the problem is defining reference ranges. It's really very different depending on how you measure it, and the link [from] the appropriate reference range to clinical impact on the patient is really, the evidence around that is very poor.*

As these examples illustrate, the argument that decisions around what constitutes appropriate testing are ultimately "judgment calls" worked to contest the notion that disinvestment policy should be guided by SR evidence in cases where evidence of effectiveness is deemed to be "complex". In these extracts, the criteria by which "appropriate testing" may be measured in evidence are complicated by the claims that "all tests are fallible", and that investment in tests as "diagnostics" or as "screening" represents a "philosophical" decision.

The final argument outlined above—that inconsistent reference ranges for deficiency cloud analysis of B_{12} test efficacy—served as the most central claim against using a SR to guide disinvestment in this area. The argument that you "cannot systematically review studies of B_{12} testing if reference ranges are inconsistent" functioned, not only to undermine the outcomes of the review presented, but to reframe the central problem and corresponding disinvestment solutions.

Reframing the Terms of the Problem

The problem is inconsistent reference ranges, so the answer is more research. By arguing that B_{12} test usage is not perfect, but "appropriate," and that any problems identified rest with inconsistent definitions of deficiency, participants contested the terms of the problem driving the SR. In turn, they contested the applicability of its outcomes in framing future policy. While participants attested to the value of systematic reviewing for disinvestment in the case of pathology items whose reference ranges and clinical expression are clear, they rejected the method in this more "complicated" case.

P2: *Where you have a test of a very reliable reference range that everybody agrees on and is appropriately linked to the clinical expression of the problem, yes, systematic review in the way you've talked about is very reliable. In this situation the thing that is complicating it is the confusion about what is the reference range of B_{12}, so applying that tool to B_{12} measurement. . . does not work because of the complexity around and the uncertainty.*

Reframing the problem as one of "reference range uncertainty" (rather than "inappropriate usage" or an "ineffective test") served to position further primary research as an appropriate next step within a disinvestment program. Significantly, the argument that such research should be undertaken within laboratories (rather than universities or government departments) served both to position pathologists as the appropriate arbiters of a disinvestment agenda and, potentially, to justify increased expenditure in this area.

P2: *One of the real concerns we have about where funding is going for pathology . . . is it's trying to strip away the professional and scientific basis of this on the grounds that, saying that this is a lousy test, just stop doing it, we're not going to fund these any more . . . and what that will do is take people like the ones sitting around this table out of the process. . . .I would say that it is not NHMRC research, it is appropriate continued research in any professionally high standard, high quality practice that's public or private . . . but it has to be recognized, the importance of that to the Australian health economy [so] that the pathology lab remuneration processes include a provision for that activity to continue.*

This argument is significant in that it supports a disinvestment path that, unlike systematic reviewing, rests upon stakeholder-defined criteria of "quality" and "standards".

The problem is folate, so we should "split the item" and "educate the doctors". As outlined earlier, a second arm of participants' arguments against disinvestment from the combined B_{12}/folate item number rested upon the notion that only folate testing is "really a problem" because mandatory fortification has significantly reduced the identification of deficiency. This argument worked, at one level, to offer an opportunity for disinvestment that pathologists described as being legitimate and evidence-based.

P3: *I think discrimination post fortification is probably the issue, in terms of folate at least that it's not really required any more. . .*

P1: *I think it could boil down to that . . . we recognize that with folate fortification, the scene has completely changed and I*

guess it's really as simple as getting the IT systems in surgeries to reflect that . . . if there's a button they press which gives a combined test, if that was de-linked it would be much more sensible.

While offering a potentially useful disinvestment option, this suggestion explicitly attributes responsibility for over-expenditure on folate tests to the clinicians who request them (not pathologists). This suggestion does not attend to the claim (presented by one participant but roundly rejected by others) that private laboratories may use the ambiguity inherent in the combined B_{12}/folate item number to test *red cell* folate (for which the laboratory receives a rebate) rather than *serum* folate (for which they are not paid).

P8: *The only reason private labs do red cell folate is because there's a rebate for it. According to the wording of the item number you do not get paid for serum folate. . .*

Moreover, the depiction of folate testing as an appropriate candidate for restriction on the grounds of fortification appears to set a very high burden of proof for disinvestment in pathology.

P2: *Maybe we are saying folate should not be done and B_{12} should be, unless you have [B_{12}] fortified food.*

Ultimately, pathologists' reframing of the "B_{12}/folate problem" as a problem of "folate alone" deflects attention from the complexity of B_{12} testing. Instead, it makes way for a solution that accords with pathologists' own interpretation of appropriate evidence for disinvestment.

DISCUSSION

Tests for B_{12} and folate are legacy items; if subjected to contemporary HTA processes as *new* or *emerging* "investment" items, the evidence base is such that it is questionable as to whether or not they would receive a favorable funding decision. This research points to the added complexity inherent in "disinvestment" decision making. Due to the distinct parallels between investment and disinvestment decision making—focusing as they both do on safety, effectiveness and cost-effectiveness—it seems logical that disinvestment be situated within the HTA/SR paradigm. The process outlined here makes explicit the central role and contribution of SR to disinvestment discussions with clinical stakeholders, in line with evidence-informed notions of robust, defensible and sustainable health policy. However, findings from this study suggest there are some limitations to this approach, grounded primarily in stakeholder's contest of the evidence and its framing.

While an evidence-informed deliberative stakeholder engagement may contribute unique solutions to a disinvestment discussion that may not have been previously considered, there are likely instances when a full SR may be of uncertain value to disinvestment initiatives, including: when there is lack of a contemporary evidence base reflecting current practice; research questions that are more appropriately addressed using existing population-based data sources (e.g., clinical registries); genuine orphan procedures with no comparator technology; co-dependent technologies; or when the salient questions relate to how a technology is applied rather than the technology itself. The manner in which these stakeholders re-framed this problem as one of "reference ranges" or "folate alone post-fortification" suggests the importance of involving stakeholders (particularly clinical) in framing the review that seeks to inform a disinvestment question. Also, their input suggests that disinvestment might look different in relation to different questions (in this case primary research may be a necessary clarifier in a disinvestment program).

Additionally, the selection of the most appropriate synthesis products may best be done in consultation with a range of stakeholders, including clinicians, to ensure methods fit for purpose. Multiple strategies for synthesizing an evidence-base exist, including (but not limited to) SRs, meta-analyses, realist synthesis (18) or rapid reviews. The appropriateness of each "synthesis product" for informing policy varies based largely upon the types of questions being asked of the evidence base: how has this problem arisen; how does the magnitude of this problem compare with other problems; what are the impacts of an intervention across specific population sub-groups; what is it about this intervention that makes it ineffective in this situation; or, how cost-effective are alternative policy or clinical approaches (13)?

Common to each synthesis product ought to be the opportunity to articulate where gaps in knowledge exist and how they might be addressed. The present analysis suggests that early, broad-based engagement with clinical stakeholders may have assisted in guiding the SR toward alternative gaps (e.g., reference range inconsistency), and identifying requisite primary research processes for addressing them. However, HTA structures in various countries may not have capacity to fill these identified research gaps. For example, funding may not be available to carry out additional primary research to clarify reference range uncertainty.

Within programs of investment/disinvestment, some jurisdictions are beginning to move toward a more consultative approach of review co-development, working alongside sponsors and stakeholders to create policy-pragmatic research. Australia is trialing SR production based on a co-developed Decision Analytic Protocol (DAP) between the review sponsor and the Medical Services Advisory Committee (MSAC), which guides the direction of a subsequent HTA incorporating a SR and economic/financial modeling. While this might flag a research gap, it does not fill it (though the review might shift in direction, if feasible). Other approaches such as "coverage with evidence development" or the "future research needs" program of the Agency for Healthcare Research and Quality could contribute important knowledge here.

While these co-produced approaches hold genuine appeal for disinvestment programs, they need to be managed carefully. Three key characteristics of disinvestment-focused initiatives

pose challenges to this process. First, disinvestment risks continued perception that it is a negative, rationing approach, rather than one espousing quality and safety. Second, co-developed disinvestment reviews may be at risk of being redirected toward ancillary issues, while neglecting critically important questions related to technology effectiveness. Finally, information asymmetry between clinically engaged experts and policy or decision makers may preclude the truly collaborative, informed and technical discussions required to generate genuine change. Overcoming these challenges will require innovative approaches to co-produced evidence syntheses, broad-based stakeholder engagements and a sustained commitment from clinicians and policy makers alike.

A stakeholder-engaged co-produced disinvestment process for building evidence synthesis products may need to consider from the outset the terms of the review; what constitutes evidence in this context; the advantages and disadvantages of different forms of evidence; and how a synthesis should be structured. Expectations on clinicians are high, as they are charged with the responsibility of representing their patients, their profession and the population as a whole. It is, therefore, critical to identify ways of supporting clinician partners in negotiating these tensions.

CONCLUSION

While engaging vested stakeholders in processes such as this poses significant challenges, if disinvestment is to remain focused on improving the quality and safety of health systems it will continue to require clinical stakeholders to inform, guide, and invest in change.

Systematic reviews and evidence syntheses are important components of quality and safety, and for informing health policy decisions, including those related to disinvestment. Trends toward co-development of SRs for informing investment decisions suggest that there may be value in similar approaches for informing disinvestment decisions, although avenues to fund identified evidence-gaps will be required.

CONTACT INFORMATION

Amber M. Watt, MPH, Senior Research Officer, Cameron D. Willis, PhD, Postdoctoral Research Fellow, Katherine Hodgetts, PhD, Research Fellow, School of Population Health and Clinical Practice, The University of Adelaide, Adelaide, South Australia, Australia

Adam G. Elshaug, MPH, PhD, Sidney Sax Fellow, School of Population Health and Clinical Practice, The University of Adelaide, Adelaide, South Australia, Australia; Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts

Janet E. Hiller, MPH, PhD, Associate Dean (Research), Professor of Public Health, Faculty of Health Sciences, Australian Catholic University, Fitzroy, Victoria, Australia; Adjunct Professor, School of Population Health and Clinical Practice, The University of Adelaide, Adelaide, South Australia, Australia

CONFLICTS OF INTEREST

Adam Elshaug has acted as consultant to the Australian Government, Department of Health and Ageing (as has Prof Hiller) and the Canadian Agency for Drugs and Technologies in Health (CADTH). The other authors report they have no potential conflicts of interest.

REFERENCES

1. Australian Government Department of Health and Ageing. Proposal for changes to the Medical Services Advisory Committee (MSAC) processes for applications for public funding. Canberra: Australian Government Department of Health and Ageing; 2010.
2. Bayram CH, Britt H, Miller G, Valenti L. *Evidence-practice gap in GP pathology test ordering: A comparison of BEACH pathology data and recommended testing*. Sydney: The University of Sydney, School of Public Health; 2009.
3. bpac nz. *Laboratory investigation of tiredness*. Dunedin: bpac nz, 2006.
4. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol.* 2006;3:77-101.
5. Edley N. Analysing masculinity: Interpretatative repertoires, subject positions and ideological dilemmas . In: Wetherell M, Taylor S, Yates SJ, eds. *Discourse as data: A guide to analysis*. London: Sage and the Open University; 2001:189-228.
6. Elshaug AG, Hiller JE, Moss JR. Exploring policy-makers' perspectives on disinvestment from ineffective healthcare practices. *Int J Technol Assess Health Care.* 2008;24:1-9.
7. Elshaug AG, Hiller JE, Tunis SR, Moss JR. Challenges in Australian policy processes for disinvestment from existing, ineffective health care practices. *Aust New Zealand Health Policy.* 2007;4:23.
8. Elshaug AG, Moss JR, Littlejohns P, et al. Identifying existing health care services that do not provide value for money. *Med J Aust.* 2009;190:269-273.
9. Gerdvilaite J, Nachtnebel A. *Disinvestment: Overview of disinvestment experiences and challenges in selected countries*. HTA-Projektbericht, Ludwig Boltzmann Institut für Health Technology Assessment; 2011.
10. Guidelines and Protocols Advisory Committee. *B12 deficiency - investigation & management of vitamin B12 and folate deficiency*. Vancouver, BC: Guidelines and Protocols Advisory Committee, 2006.
11. Lin V, Gibson B. Introduction. In: Lin V, Gibson B, eds. *Evidence-based health policy. Problems and possibilities*. Melbourne: Oxford University Press; 2003.
12. Lomas J, Culyer T, McCutcheon C, McAuley L, Law S. Conceptualizing and combining evidence for health system guidance. Ottawa: Canadian Health Services Research Foundation; 2005.
13. Mays N, Pope C, Popay J. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *J Health Serv Res Policy*. 2005;10:16-20.
14. Medical Benefits Reviews Task Group. *Review of the funding arrangements for pathology services*. Canberra: Department of Health and Ageing; 2011.
15. Medicare Australia. *MBS items statistics reports*. Canberra: Department of Health and Ageing; 2011.
16. Menon D, Stafinski T. Health Technology Assessment in Canada: 20 years strong? *Value Health.* 2000;12:14-19.

17. Ontario Association of Medical Laboratories. *Guidelines for folate testing (CLP-026)*. New York: Ontario Association of Medical Laboratories; 2009.
18. Pawson R. Evidence-based policy: The promise of 'realist synthesis'. *Evaluation.* 2002;8:340-358.
19. Pearson SD, Rawlins MD. Quality, innovation, and value for money. NICE and the British National Health Service. *JAMA*. 2005;294:2618-2622.
20. Royal College of Pathologists of Australasia. *RCPA Manual*. Sydney: Royal College of Pathologists of Australasia; 2009.
21. Sampietro-Colom L, Ausa J, Briones E, Gol J, AuNETS Group. History of health technology assessment: Spain. *Int J Technol Assess Health Care*. 2009;25:163-173.
22. Willis CD, Elshaug AG, Milverton JL, et al. Diagnostic performance of serum cobalamin tests: A systematic review and meta-analysis. *Pathology*. 2011;43:472-481.