

BMJ Open Development and validation of a screening tool for the identification of inappropriate transthoracic echocardiograms

Ricardo Fonseca,¹ Faraz Pathan,¹ Thomas H Marwick²

To cite: Fonseca R, Pathan F, Marwick TH. Development and validation of a screening tool for the identification of inappropriate transthoracic echocardiograms. *BMJ Open* 2016;**6**:e012702. doi:10.1136/bmjopen-2016-012702

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2016-012702>).

RF and FP should be regarded as joint first authors.

Received 18 May 2016
Revised 23 July 2016
Accepted 9 September 2016



CrossMark

¹Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia

²Baker IDI Heart and Diabetes Institute, Melbourne, Tasmania, Australia

Correspondence to

Dr Thomas H Marwick; tom.marwick@bakeridi.edu.au

ABSTRACT

Objective: We sought whether simple clinical markers could be used in a questionnaire for recognition of inappropriate (or rarely appropriate, RA) tests at point-of-service. Most applications of appropriateness criteria (AC) for transthoracic echocardiogram (TTE) have been at the point of order, but a simple means of identifying RA tests in an audit process would be of value.

Design, setting and participants: The study was performed in 2 major hospitals in Tasmania. 2 reviewers created a questionnaire based on 4 questions most commonly associated with RA (suspected endocarditis with no positive blood cultures or new murmur, lack of cardiovascular symptoms or no change in clinical status or cardiac examination, routine surveillance and previous TTE within a year) in a derivation cohort of 814 patients. This was prospectively applied to 499 TTEs to calculate sensitivity and specificity for prediction of RA, and validated in the external group (n=880).

Results: Of 499 prospective TTEs, the questionnaire selected 18% requests as being potentially RA. As 7.4% were actually RA (κ 89%), the sensitivity and specificity of the questionnaire were 84% and 87%, respectively. In the external validation cohort, the model found 11% requests needed to be screened for appropriateness with a sensitivity and specificity of 80% and 95%.

Conclusions: A questionnaire based on 4 questions detects a high proportion of RA TTE, and could be used for audit.

INTRODUCTION

Investigations constitute an important component of resource utilisation and waste in medicine. The appropriateness criteria (AC) for transthoracic echocardiography (TTE)¹ have become widely adopted in the USA, and seek to control resource usage, reduce variability of practice and to improve decision-making and patient care.² Evaluations of appropriate use have exposed potential targets for improvement,^{3 4} but

Strengths and limitations of this study

- Four binary questions encapsulate characteristics of rarely appropriate (RA) tests according to the appropriate use criteria for echocardiography.
- The questionnaire, applied to the transthoracic echocardiogram requests, selected around one in five requests as being potentially RA.
- Two or more affirmative answers had a high sensitivity and specificity to discriminate RA tests.
- It is a feasible tool which can be used at the point of service to screen for inappropriate tests with a low impact on the workflow.
- The use of this approach requires review of medical records to adjudicate appropriateness when inadequate information is provided on the request form.

despite attempts to decrease rarely appropriate (RA) testing at the point of order, there is limited evidence of a decline in the number of tests,^{5–9} and in some cases, results are contradictory.^{4 10} AC are probably not responsible for the reduction in imaging over the past decade;¹¹ the trend of appropriate and RA tests has not improved during this time.¹² This lack of clear improvement in the rate of RA tests¹² contrasts with the heightened level of awareness on RA test use.^{2 5 13} A recent review noted that guidelines for quality in cardiovascular imaging advocate implementation of AC, though there are limited effective methods to reduce the rate of inappropriate testing. They contend that part of the complexity of implementing AC may be unfamiliarity with the classification and the time required to review each imaging request with the guidelines.¹⁴

Various processes have been devised to police RA requests by screening at point-of-order, for example, based on the use of radiology benefit managers or software integrated with the ordering process.^{15–17} An

alternative approach might be laboratory-based,⁶ but matching patient details against >100 AC is impractical and inefficient. In this study, we sought to determine whether a simple point-of-service questionnaire (PSQ) based on the most common RA characteristics according to the 2011 AC for TTE^{1 18} could facilitate recognition of these tests in the echocardiography laboratory.

METHODS

We sought to develop the PSQ and then perform a diagnostic accuracy analysis.¹⁹ The model was developed in three steps (figure 1). Eligible requests were selected from two separate hospitals and at the times specified in steps 1–3. TTE requests were excluded for analysis if the patient was <18 years, or if classification was not possible because of insufficient clinical documentation.

The study design pertained to appropriate selection of tests already ordered by the patient's physician. We elected not to discuss the uncertainties about appropriate use with these patients. Nonetheless, we have discussed appropriate use with patient representatives in meetings about medical expenditure in Australia. Because of the copayments associated with outpatient echocardiography in this country, this is perceived as a very important topic.

Our target condition was to determine 'inappropriate' (also known as 'rarely appropriate', RA) TTE requests at the point of service as adjudged by the reference standard with the 'Appropriateness Criteria for echocardiography (AC)'.¹ We developed then assessed the index test (PSQ). We compared the accuracy of our index PSQ model with the reference AC standard by evaluating sensitivity, specificity, positive and negative predictive value for each of the responses as well for each of the possible cut points: one affirmative answer versus two affirmative answers versus three affirmative answers.

1. Derivation of model: The most common causes of inappropriate tests described by the 2011 AC for TTE were identified from a retrospective group of 814 TTE requests, at a teaching hospital. After this analysis, four questions which summarise those characteristics were identified. The 'PSQ' comprised binary questions based on the characteristics of RA tests in our derivation cohort and also accounted for published characteristics seen in the choosing wisely programme¹⁴ and published research.¹⁸
2. Internal validation: We then tested these questions in an internal validation cohort to ascertain their ability to identify RA requests as judged against the gold standard (AC code for a specific request). The four most common characteristics for RA tests (PSQ) were applied prospectively to all the requests (n=499) for a TTE at the same tertiary referral hospital, between March and May 2015.
3. External validation: The PSQ was applied to a cohort of 880 requests at a regional referral hospital between May and August 2015.

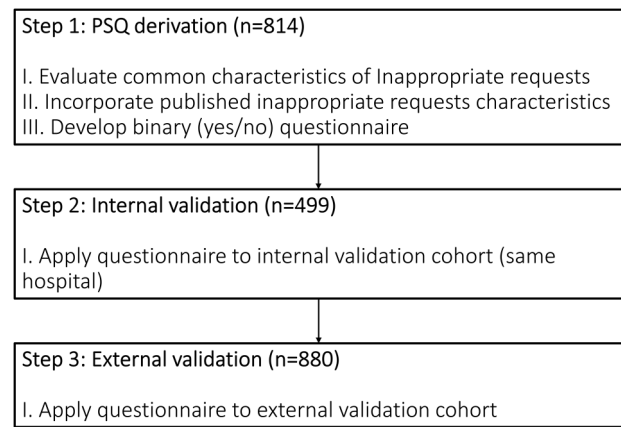


Figure 1 Design of the study.

Patient demographic information, inpatient/outpatient distribution, referring physician (cardiologist or not) and the indication for the study were determined from the request form. Investigators reviewed the digital medical record to capture any additional information, especially when confronted with inadequate information. The time required to access additional information was recorded whenever such an action was necessary and the result was averaged for all such requests.

For each of the steps, a general physician (RF) and a cardiologist (FP) independently recorded the results for each of the questions in the PSQ. Appropriateness of requests was scored by the same observers, independently of the PSQ evaluation, using the 2011 AC.¹ Each study was scored as appropriate, RA (previously described as 'inappropriate') or maybe appropriate (previously 'uncertain'). If the main indication was not listed in the AC, investigators were asked to select 'Not classifiable'.² When there was disagreement, a consensus between reviewers was reached. If no consensus was attained, a third investigator (THM) reviewed the data and determined AC score. This AC score served as the 'Gold Standard' by which the PSQ was assessed. Where a repeat study was performed to guide management (eg, repeat TTE to evaluate reverse remodelling after 3–6 months of medical therapy), despite two affirmative responses, this request was classified as appropriate as per the AC guidelines on studies which are used to guide management.

The sensitivity, specificity, positive likelihood ratio and OR were used to define the PSQ accuracy for the prediction of RA requests using R software (R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. 2008. <http://www.R-project.org>). The above parameters were analysed for each affirmative response and cumulatively. During individual analysis, a single affirmative response (question 1, 2, 3 or 4) was compared with 'no affirmative responses'. For the assessment of the cumulative affirmative responses, comparison was made to 'no affirmative responses' and to 'no or one affirmative response'.

Interobserver variability in the scoring process to determine level of appropriateness was defined using κ statistics.

RESULTS

Predictors of RA tests were sought in a group of 814 patients among whom 9% were RA. Our results revealed the RA requests corresponded with indications where there were no new symptoms or no change in clinical status or cardiac examination (indications 35, 53, 10, 8 of the AC) and indications for routine surveillance (indications 88, 11, 13, 40, 28). We sought to distil the underlying markers of an inappropriate request using 'routine studies' and the 'absence of a change in clinical status or new symptoms' which accounted for 88% of RA requests. Furthermore, 28% of RA tests had a TTE within the previous year. Additionally, there was a specific clinical situation that accounted for the 26% of RA tests (Indication 53: transient fever without evidence of bacteraemia or new murmur).

We identified four features associated with RA tests: evaluation in the absence of symptoms/signs of cardiac disease or no change in clinical status, routine surveillance, existence of a previous TTE within the year of the new TTE request and suspected endocarditis with no positive blood cultures or new murmur (table 1).

Based on the above, we developed a PSQ composed of four binary questions:

Q1: Was the scan requested in the absence of new cardiovascular symptoms, or change in clinical status or cardiac examination? Note that it requires symptoms to be cardiovascular (this would include transient ischaemic attacks, strokes). Pre-existing symptoms or signs such as a long-standing murmur or dyspnoea which has been evaluated and have not changed would score as a 'yes' response. Therefore, a 'yes response' (affirmative response) to question 1 means the patient does not have any new cardiovascular sign or symptom or in those with pre-existing cardiovascular illness, there has not been a worsening of their clinical status.

Q2: Is this a routine surveillance scan? This captures tests being considered for a 'periodic' evaluation since a certain period of time has elapsed. The test is not being ordered due to the anticipation of changing clinical decision-making or guiding therapy.

Q3: Has there been a previous TTE within the last year?

Q4: Is the test requested for suspected endocarditis with no positive blood cultures or new murmur?

PSQ was applied to 501 studies in the internal validation cohort at a tertiary referral hospital and to 881 TTE requests at a regional hospital (external validation cohort). Two requests within the internal validation cohort and one request in the external validation cohort were not classifiable as it was not possible to collect information to answer the questionnaire. Final analysis was made in 499 TTE requests in the internal validation

Table 1 Rarely appropriate tests found in derivation group

AC item	Count	Proportion surveillance	Routine surveillance	Lack of change in clinical status/evaluation of symptoms without other symptoms/signs of cardiac disease
35 Initial evaluation when there are no other symptoms or signs of valvular or structural heart disease	25	0.36	No	Yes
53 Transient fever without evidence of bacteraemia or a new murmur	18	0.26	No	Yes
10 Initial evaluation of ventricular function (eg, screening) with no symptoms or signs of cardiovascular disease	6	0.09	No	Yes
28 Suspected pulmonary embolism in order to establish diagnosis	4	0.06	No	No
54 Transient bacteraemia with a pathogen not typically associated with infective endocarditis and/or a documented non-endothelial source of infection	4	0.06	No	No
88 Routine surveillance (<1 year) of known cardiomyopathy without a change in clinical status or cardiac examination	3	0.04	Yes	Yes
8 Light-headedness/presyncope when there are no other symptoms or signs of cardiovascular disease	2	0.03	No	Yes
11 Routine surveillance of ventricular function with known CAD and on ou change in clinical status or cardiac examination	2	0.03	Yes	Yes
13 Routine perioperative evaluation of ventricular function with no symptoms or signs of cardiovascular disease	2	0.03	Yes	Yes
40 Routine surveillance (>1 year) of moderate or severe valvular stenosis without a change in clinical status or cardiac examination	2	0.03	Yes	Yes
38 Routine surveillance (<3 years) of mild valvular stenosis without a change in clinical status or cardiac examination	1	0.01	Yes	Yes

and 880 TTE requests in the external validation groups, respectively.

Table 2 shows study characteristics and appropriateness classification by group. The internal and external validation cohorts are well matched; however, the former group had a lower proportion of outpatients. The 10 most common RA indications in the internal validation group are described in **table 3**. Inter-rater agreement for scoring between both reviewers was high ($\kappa=89\%$).

When question 1 (no change in symptoms/no change in clinical status/no change in cardiac examination) was answered affirmatively, it had a higher OR, greater positive predictive value and positive likelihood ratio to pick up possible RA tests when compared independently with other questions. This was driven by examination of asymptomatic or stable patients (see online supplementary table S1).

In the internal validation cohort, 18% of the tests had ≥ 2 affirmative answers (two 'yes' responses). A PSQ with ≥ 2 affirmative responses had an OR 33.96 (13.61 to 84.78), sensitivity 0.84 (0.68 to 0.94) and specificity of 0.87 (0.83 to 0.90) for RA.

In the external validation group (n=880), ≥ 2 affirmative answers were 11% of total requests; it had a sensitivity of 80%, specificity of 95% and OR 83.01 (**table 4**).

The PSQ with ≥ 2 affirmative responses identified 84%, and 80% of the inappropriate tests in the internal and external validation cohorts, respectively.

Around 20% of the TTE requests provided inadequate information and it was necessary to check the digital medical records (**table 5**).

DISCUSSION

The approach proposed in this study has been to encapsulate the essence of the RA tests into four binary questions which can be used rapidly to screen for these inappropriate requests. The questions which we have used are consistent with published literature with the choosing wisely programme identifying 'Routine studies' and 'no change in signs or symptoms' as unnecessary repeat testing.¹⁴ Similarly, other authors identified that 54% of inappropriate requests had a TTE within the last year.¹⁸ Although only 28% of the inappropriate studies

Table 2 Study characteristics and appropriateness classification according to groups

	PSQ derivation cohort (n=814)	Internal validation cohort (n=499)	External validation cohort (n=880)	p Value
Age, years (median (IQR))	65.00 (52.00–76.00)	67 (55–76)	70 (58–80)	<0.01
Male, n (%)	444 (54.6)	256 (51.3)	449 (51.0)	0.96
Outpatient, n (%)	573 (70.4)	289 (57.9)	661 (75.1)	<0.01
Referred by cardiologists, n (%)	348 (42.8)	253 (50.7)	686 (78.0)	<0.01
Appropriateness score				
Appropriate, n (%)	707 (86.9)	431 (86.4)	774 (88.0)	<0.01
RA appropriate, n (%)	68 (8.4)	37 (7.4)	75 (8.5)	
May be appropriate, n (%)	13 (1.6)	18 (3.6)	7 (0.8)	
Unclassifiable, n (%)	26 (3.2)	13 (2.6)	24 (2.7)	

Data in bold typeface denote statistically significant results.

Table 3 Ten most common rarely appropriate indications in prospective internal validation cohort

Indication	#	Proportion
88 Routine surveillance (<1 year) of known cardiomyopathy without a change in clinical status or cardiac examination	7	0.19
10 Initial evaluation of ventricular function (eg, screening) with no symptoms or signs of cardiovascular disease	5	0.14
53 Transient fever without evidence of bacteraemia or a new murmur	4	0.11
60 Routine surveillance of known small pericardial effusion with no change in clinical status	3	0.08
35 Initial evaluation when there are no other symptoms or signs of valvular or structural heart disease	2	0.05
36 Re-evaluation in a patient without valvular disease on prior echocardiogram and no change in clinical status or cardiac examination	2	0.05
40 Routine surveillance (<1 year) of moderate or severe valvular stenosis without a change in clinical status or cardiac examination	2	0.05
48 Routine surveillance (<3 years after valve implantation) of prosthetic valve if no known or suspected valve dysfunction	2	0.05
74 Routine surveillance (<1 year) of HF (systolic or diastolic) when there is no change in clinical status or cardiac examination	2	0.05
79 Routine surveillance (<1 year) of implanted device without a change in clinical status or cardiac examination	2	0.05

Table 4 Diagnostic tests of the point-of-service questionnaire for rarely appropriate requests

	Sensitivity estimate (lower–upper)	Specificity estimate (lower–upper)	OR estimate (lower–upper)	Positive predictive value estimate (lower–upper)	Negative predictive value (lower–upper)	Positive likelihood ratio estimate (lower–upper)	Negative likelihood ratio estimate (lower–upper)
Prospective cohort	0.84 (0.68 to 0.94)	0.87 (0.83 to 0.90)	33.96 (13.61 to 4.78)	0.34 (0.24 to 0.44)	0.98 (0.97 to 1.00)	6.34 (4.83 to 8.34)	0.19 (0.09 to 0.39)
External retrospective validation cohort	0.80 (0.69 to 0.88)	0.95 (0.84 to 0.97)	83.3 (43.13 to 159.83)	0.62 (0.51 to 0.72)	0.98 (0.97 to 0.99)	17.41 (12.46 to 24.32)	0.21 (0.13 to 0.33)

in the derivation cohort had a TTE within the previous year, we included this question to improve the strength of our model. Finally, given the high prevalence of inappropriate endocarditis requests (26% in our derivation cohort), it was included as the final question in our model.

The primary finding of this study is that the application of the PSQ is feasible, and it identifies a high proportion of RA tests without the need to review all the TTE requests against the AC. An affirmative response to any of the questions increases the likelihood of a test being deemed RA, and when two (or more) of the questions were answered affirmatively, the chance of determining a test as inappropriate increased more than 33 times. The results were consistent in the different cohorts and scenarios (inpatients/outpatients, test referred by cardiologists/non-cardiologists). We propose a PSQ-based method for screening appropriateness (figure 2). Using this model less than one-fifth of TTE requests would need to be audited against the AC, thereby minimising interruption of workflow. Nevertheless, occasional requests such as asymptomatic severe mitral regurgitation surveillance within 1 years (AC indication 45: uncertain) or repeat echocardiography for a heart failure patient on optimal medical therapy without a change in signs or symptoms to guide therapy (AC indication 73: appropriate) could still be performed, despite two affirmative responses.

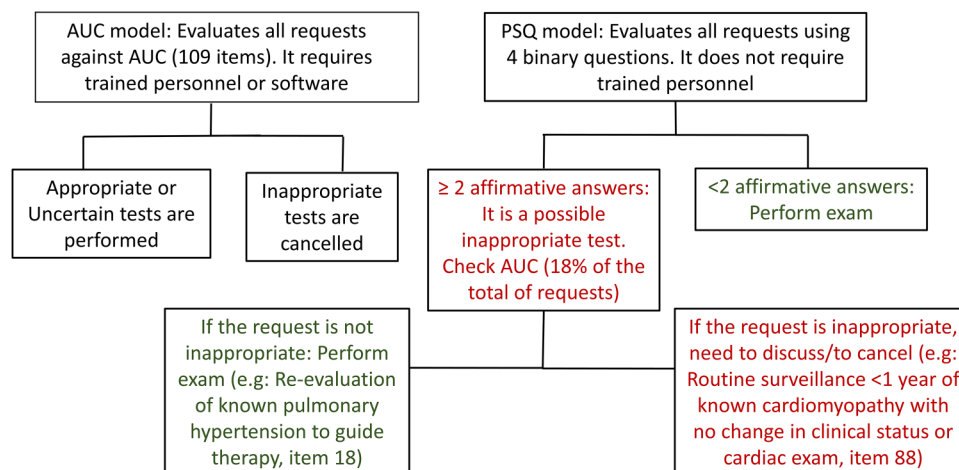
The use of radiology benefit management companies (RBM) is still an important pole in the process of performance of cardiac imaging, although one of the aims of the AC was to reduce the need of those companies.²⁰ Prior authorisation and claim denials continue to be the top challenges of the process.²¹ The results of this study show that the use of the questionnaire provides a transparent solution which can be implemented with minimal delay at point-of-service, thereby minimising the need for RBMs or other middlemen.

Several attempts have been made to improve appropriateness at the point-of-order, implementing software to control the request of inappropriate tests.^{7 17 22} However, the use of those tools at the point-of-order is susceptible to indication drift: for example, the real indication for testing may be RA, but inactive problems are appropriate. Perhaps for this reason, the AUC literature has shown little or no improvement in requesting behaviour.^{7 17} Our proposed method could be used to facilitate appropriate use audits (simplifying to 4 questions from >200 AUC, which is useful where these data are not available in electronic format), or added to the current appropriate use process at the point-of-service. At that level, the proposed approach provides a simple screening tool to flag possibly inappropriate tests at the time of scheduling.

In our study, the prevalence of RA tests varied between 7% and 9%. These values fall on the lower end of the prevalence distribution of RA tests documented in various studies.^{4 12 18 22–25} A comprehensive plan

Table 5 Differences in time when medical record (DMR) needed to be checked

	Internal validation cohort		External validation cohort	
	Yes	No	Yes	No
Needed DMR	95 (0.19)	404 (0.81)	168 (0.19)	712 (0.81)
Seconds (median (IQR))	120 (62–120)	18 (14–21)	80 (63–98)	12 (9–20)

**Figure 2** Comparison between the appropriateness criteria model and point-of-service-questionnaire model.

detailing the management of inappropriate requests found at the point of service is lacking. At the very least, a clear and effective communication strategy needs to be in place to inform a discussion before a 'rarely appropriate' test is scheduled.

This questionnaire should act as a prompt to refer to the AC rather than an absolute assessment of appropriateness. Education has resulted in increased awareness of the AC without a significant change in clinical practice.¹² It is intuitive to entrust point-of-care policing regarding the appropriateness of an investigation to those with the greatest experience. However, such a policy may result in interruption of workflow, delays and a greater burden on already busy echocardiography units. Our results show that the use of four simple binary questions identifies RA tests with a high sensitivity and specificity. The questionnaire identifies potential RA requests which can then be confirmed by reference to the AC (available online, inprint or as a mobile phone applications) or by discussing the case with a cardiologist at the echocardiographic laboratory or with the referring physician.

The use of the questionnaire may face some challenges. First, the need to review medical records to adjudicate appropriateness when inadequate information was provided on request forms is a potential limitation. However, corroboration of clinical history is a common clinical practice in all imaging units as it is mandatory to establish the question being asked of an investigation and is essential to implementing a Bayesian approach to reporting. This raises a second issue which is quality control of echocardiography requests. We identified

20% of requests as inadequate requiring further corroboration of clinical history. This result was similar to recent published data which found in a review of 1303 requests that 26.2% were inadequate to determine adherence to AC.²⁶ The study concluded that the top three reasons for inadequacy were failure to report change in clinical status or cardiac examination, date of prior echocardiogram and type and severity of valvular lesion. Though the last two would be easily accessible in any echocardiographic laboratory (assuming the previous studies were performed within that laboratory), the first failure is critical in determining appropriateness. Clearly, it follows that access to electronic medical records is a necessary and essential component of an echocardiographic laboratory's workflow.

Second, the possible difference in referral patterns between hospitals and private labs may impact on the positive predictive value of our questionnaire. The private/public divide varies tremendously across countries and to date has not been assessed in regard to compliance with AC. Though our criteria address the issue of awareness and simplicity perhaps the greatest challenge facing an overburdened medical infrastructure is the systemic dependence of investigations.

Furthermore, although our rate of inappropriate use is lower than reported in other institution, and could be seen as a limitation, we overcame this issue by performing an analysis of over 800 requests in the derivation process. We also validated this questionnaire in over 1200 patients (internal and external validation cohorts).

Finally, previous researchers have sought to differentiate the appropriateness of a study from its clinical utility

arguing an RA test does not necessarily mean a clinically useless one nor does an appropriate request always correspond with a useful one.^{18 25} Ward demonstrated that 17% of inappropriate tests had 'new important TTE abnormalities' and Matulevicius showed that 21.7% of RA tests led to an active change in management. *Thus, while the identification of inappropriate tests is a step on the path to improving quality and appropriateness in cardiovascular imaging, decision-making has to be informed by individual characteristics.*

How we handle RA requests will ultimately have financial and clinical implications. The proportion of inappropriate studies varies between 7% and 23%.^{12 14} Our study shows an inappropriate rate of 7.4–8.5%. Experience from elsewhere in Australia has demonstrated inappropriate rate of 20%.²⁷ In 2015, the total Medicare reimbursement for TTE was ~A\$186.0 million.²⁸ Assuming a rate of inappropriate echocardiography between 7% and 20%, the cost to the Australian health system of inappropriate TTE would be between AU\$13.0 and 37.2 million. The healthcare costs are clearly proportionate to the use of TTE and prevalence of appropriate use.

A mandatory AC score (appropriate or otherwise) or point-of-service score tied to funding would enhance compliance with AC and enable continuous auditing of resource usage. There are >100 categories of appropriateness and the incremental workflow issues are prohibitive. We proposed an alternative approach where response to the PSQ serves as a less cumbersome beacon of appropriateness.

We have demonstrated that ≥ 2 affirmative answers at a simple PSQ detect a high proportion of RA tests. This approach can be used as a red flag for inappropriate examinations and a prompt to further discussion about the suitability for testing in individual patients. We propose this PSQ can be a quality control tool that captures the majority of inappropriate use, in the absence of the infrastructure that supports AC in North America, and a simple marker for departmental and regional audits.

Contributors THM and RF designed the study. RF and FP were responsible for data collection and data analysis under supervision of THM. RF and FP prepared the manuscript draft. THM edited and approved the final manuscript. All the authors had full access to all of the data (including statistical reports and tables) in the study.

Funding RF is supported by a scholarship from the Farrell Foundation, Hobart, Australia.

Competing interests None declared.

Ethics approval This study was approved by the Tasmanian Human Research Ethics Committee, reference number: H0014017.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

- Douglas PS, Garcia MJ, Haines DE, *et al.* ACCF/AHA/ASNC/HFSA/HRS/SCAI/SCCM/SCCT/SCMR 2011 appropriate use criteria for echocardiography. A report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Society of Echocardiography, American Heart Association, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Critical Care Medicine, Society of Cardiovascular Computed Tomography, and Society for Cardiovascular Magnetic Resonance Endorsed by the American College of Chest Physicians. *J Am Coll Cardiol* 2011;57:1126–66.
- Hendel RC, Patel MR, Allen JM, *et al.* Appropriate use of cardiovascular technology: 2013 ACCF appropriate use criteria methodology update: a report of the American College of Cardiology Foundation appropriate use criteria task force. *J Am Coll Cardiol* 2013;61:1305–17.
- Ward RP, Krauss D, Mansour IN, *et al.* Comparison of the clinical application of the American College of Cardiology Foundation/American Society of Echocardiography Appropriateness Criteria for outpatient transthoracic echocardiography in academic and community practice settings. *J Am Soc Echocardiogr* 2009;22:1375–81.
- Bhatia RS, Milford CE, Picard MH, *et al.* An educational intervention reduces the rate of inappropriate echocardiograms on an inpatient medical service. *JACC Cardiovasc Imaging* 2013;6:545–55.
- Lin GA, Harris IS. Can appropriate use criteria lead to appropriate care? *Circ Cardiovasc Qual Outcomes* 2015;8:4–5.
- Fonseca R, Marwick TH. How I do it: judging appropriateness for TTE and TEE. *Cardiovasc Ultrasound* 2014;12:22.
- Goldzweig CL, Orshansky G, Paige NM, *et al.* Electronic health record-based interventions for improving appropriate diagnostic imaging: a systematic review and meta-analysis. *Ann Intern Med* 2015;162:557–65.
- Levitt K, Shojania KG, Bhatia RS. Point-of-care decision support for reducing inappropriate test use: easier said than done. *BMJ Qual Saf* 2016;25:6–8.
- Boggan JC, Schulte RD, Donahue M, *et al.* Guideline-based decision support has a small, non-sustained effect on transthoracic echocardiography ordering frequency. *BMJ Qual Saf* 2016;25:57–62.
- Gibbons RJ, Miller TD, Hodge D, *et al.* Application of appropriateness criteria to stress single-photon emission computed tomography sestamibi studies and stress echocardiograms in an academic medical center. *J Am Coll Cardiol* 2008;51:1283–9.
- Medicare Payment Advisory Commission (MedPAC). *Report to the Congress: Medicare Payment Policy*. 2013. http://www.medpac.gov/documents/reports/mar13_entirereport.pdf?sfvrsn=0 (accessed 29 Jun 2015).
- Fonseca R, Negishi K, Otahal P, *et al.* Temporal changes in appropriateness of cardiac imaging. *J Am Coll Cardiol* 2015;65:763–73.
- Intersocietal Accreditation Commission (IAC) Echocardiography (ICAEL). IAC Standards and Guidelines for Adult Echocardiography Accreditation. 2013. <http://www.intersocietal.org/echo/standards/IACAAdultEchoStandardsJuly2014.pdf> (accessed 29 Jun 2015).
- Bhattacharyya S, Lloyd G. Improving appropriateness and quality in cardiovascular imaging: a review of the evidence. *Circ Cardiovasc Imaging* 2015;8:pii: e003988.
- Willens HJ, Nelson K, Hendel RC. Appropriate use criteria for stress echocardiography: impact of updated criteria on appropriateness ratings, correlation with pre-authorization guidelines, and effect of temporal trends and an educational initiative on utilization. *JACC Cardiovasc Imaging* 2013;6:297–309.
- Hendel RC, Cerqueira M, Douglas PS, *et al.* A multicenter assessment of the use of single-photon emission computed tomography myocardial perfusion imaging with appropriateness criteria. *J Am Coll Cardiol* 2010;55:156–62.
- Lin FY, Dunning AM, Narula J, *et al.* Impact of an automated multimodality point-of-order decision support tool on rates of appropriate testing and clinical decision making for individuals with suspected coronary artery disease: a prospective multicenter study. *J Am Coll Cardiol* 2013;62:308–16.
- Ward RP, Mansour IN, Lemieux N, *et al.* Prospective evaluation of the clinical application of the American College of Cardiology Foundation/American Society of Echocardiography Appropriateness Criteria for Transthoracic Echocardiography. *JACC Cardiovasc Imaging* 2008;1:663–71.



19. Bossuyt PM, Reitsma JB, Bruns DE, *et al.* STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;351:h5527.
20. Dehmer GJ, Patel MR. The use of appropriate use criteria are increasing, but what are their effects on medical care? *Circulation* 2015;132:4–6.
21. Shor R. Using appropriate use criteria to address pre-authorization. *J Am Coll Cardiol* 2015;66:1300–2.
22. Bhave NM, Mansour IN, Veronesi F, *et al.* Use of a web-based application of the American College of Cardiology Foundation/American Society of Echocardiography Appropriateness Use Criteria for Transthoracic Echocardiography: a pilot study. *J Am Soc Echocardiogr* 2011;24:271–6.
23. Bhatia RS, Carne DM, Picard MH, *et al.* Comparison of the 2007 and 2011 appropriate use criteria for transthoracic echocardiography in various clinical settings. *J Am Soc Echocardiogr* 2012;25:1162–9.
24. Kirkpatrick JN, Ky B, Rahmouni HW, *et al.* Application of appropriateness criteria in outpatient transthoracic echocardiography. *J Am Soc Echocardiogr* 2009;22:53–9.
25. Matulevicius SA, Rohatgi A, Das SR, *et al.* Appropriate use and clinical impact of transthoracic echocardiography. *JAMA Intern Med* 2013;173:1600–7.
26. Banihashemi B, Maftoon K, Chow BJ, *et al.* Limitations of free-form-text diagnostic requisitions as a tool for evaluating adherence to appropriate use criteria for transthoracic echocardiography. *Cardiovasc Ultrasoun* 2015;13:4.
27. Al-Kaisey A, Jones E, Nadurata V, *et al.* Appropriate use of echocardiography in an Australian regional centre. *Intern Med J* 2015;45:1128–33.
28. Australian Government Department of Human Services. *Medicare Australia Statistics, Medicare Item Reports*. https://www.medicareaustralia.gov.au/statistics/mbs_item.shtml (accessed July 2016).