"Appropriate Use" of Cardiac Imaging Principle: Is "Appropriate Use" of Diagnostic Tests Similar to Their Clinical Usefulness?

Randomized clinical trials of diagnostic tests should be encouraged. These trials need to focus on specific clinical scenarios and field experts may be consulted on how to define them.
John P. A. Ioannidis

INTRODUCTION

If abstract reasoning is used, why not think that if the diagnosis of a cardiovascular disease is obtained with an imaging test (screening) in an asymptomatic patient, treatment will be started earlier and the course of the disease will change?

Moreover, we all know that cardiovascular disease (CVD) is the leading cause of death in most countries worldwide, at least when a person reaches adulthood, so physicians and the general population have great interest in identifying individuals at risk of CVD to implement preventive measures.

Perhaps risk prediction beyond the assessment of the so-called "risk factors" could be improved with other clinical diagnostic tests, including the new cardiovascular imaging studies, although there is still no evidence that changes regarding management occur or long-term outcomes improve.

Transthoracic echocardiography (TTE) is the most popular cardiac imaging technique among both physicians and patients, representing about half of all the different types of cardiac imaging studies performed, thus constituting a widely used clinical tool for the diagnosis and management of cardiovascular disease.

All doctors have gone through the personal experience of discovering that an echocardiogram in asymptomatic individuals may reveal incidental findings such as hypertrophic cardiomyopathy, valve disease (mainly aortic valve disease) or other rarer diseases, and also that some of these hidden diseases may be the cause of unexplained cardiac deaths among adults and athletes.

We could mention that the guideline of the American College of Cardiology Foundation / American Heart Association classifies resting echocardiographic screening with a Class IIb indication (benefit slightly greater than risk) to detect hypertrophy and left ventricular dysfunction in asymptomatic adults with hypertension , but classifies it as Class III (no benefit) in asymptomatic adults without hypertension. (1)

In fact, because of its safe and easy implementation, most echocardiograms are ordered by primary care physicians rather than by cardiologists. Therefore, since its use has grown exponentially in the last years, echocardiography is already being used *de facto* as screening of asymptomatic subjects. In Medicare (USA), the use of echocardiography grew 90% in 10 years (1999-2008). (2)

The linear abstract reasoning does not take into consideration the complex context with multiple variables that can affect the final outcome of a patient who undergoes TTE. The controlled clinical experience of randomized trials can answer this enigma; let us then look at the literature.

ECHOCARDIOGRAPHIC SCREENING OF THE GENERAL POPULATION AND LONG-TERM SURVIVAL

Lindekleiv et al. have recently published the results of the Tromsø cohort, a population-based study performed in Norway. (3) During the fourth survey (1994-1995), all the inhabitants aged 25 years or older were invited to participate. Among the 21159 eligible subjects, 77% attended the first visit. All individuals aged 55 to 74 years and a randomized sample of other ages was invited to a second visit where they underwent more extensive screening. The study population consisted of 6861 subjects who attended the second visit and gave informed consent for the study.

This group of 6861 persons was randomly assigned to a type 1 or 2 test. Generally, this design essentially randomized non-selected middle-aged subjects to a screening strategy with or without echocardiography. This design was not initially intended as a clinical trial with an a priori hypothesis concerning the usefulness of echocardiographic screening in reducing long-term mortality. Actually, they were randomized to avoid a selection bias, since only one of the two test branches included the use of echocardiograhy, due to the inability of conducting it in all subjects. In the end, patients were randomized to echocardiography or control, with no difference between the two branches regarding the rest of the exams.

Average age was 60 years (SD± 10 years) and the percentage per gender was practically divided in halves. Mean blood pressure was 145/83 mm Hg, almost 60% of subjects had history of hypertension and almost 32% smoked. However, the use of antihyper-

tensive medication and statins was less than 14% and 2%, respectively.

In the group of 3272 participants assigned to echocardiography, 362 (11%) had findings that justified referral to a cardiologist and 290 (8.9%) were evaluated as a result of screening. Significant incidental findings were valve disease present in 3.3% of subjects, and occasionally hypertrophic cardiomyopathy, ventricular dysfunction, wall motion abnormalities, myxoma, and other conditions.

Despite cardiological referral due to these diagnostic findings, mortality during the 15-year follow-up period did not differ between patients with or without echocardiographic screening (26.9% vs 27.6%, HR 0.97 95% CI 0.89-1.06). There were neither differences in cardiac death, sudden death and incidence of myocardial infarction [(HR 0.95 (0.83 - 1.08), or stroke (HR 1.02 (0.87 - 1.19)].

No significant differences in the predefined subgroups of patients with hypertension (34% of the population, contrary to guideline indications) (1), diabetes, family history of early myocardial infarction, or risk of fatal cardiovascular disease was observed at 10 years.

Findings of this cohort of 6861 middle-aged subjects, which was very representative of the population and with a mortality rate similar to that of the general population across Norway, provides for the first time firm evidence that echocardiography as screening of structural and valvular heart disease affords no benefit in mortality, or in myocardial infarction or stroke risk. It has neither benefit in higher risk subgroups such as hypertensive or diabetic patients, and those with a family history of myocardial infarction or increased risk of fatal cardiovascular disease.

These findings reinforce the idea that echocardiographic screening of structural heart disease in asymptomatic patients offers no benefit, and therefore, should not be performed. This indication should be taken into account today, even more so when pocket-sized echocardiography may become a routine tool in clinical practice.

Although echocardiography is noninvasive and does not emit radiation, unjustified use would not be exempt of some risks. Due to incidental findings, additional studies may lead to anxiety, psychological damage and possible complications owing to further research, without any clinical benefit. In contrast, a normal echocardiogram may provide a false reassurance inducing the patient not to perform studies that would be indicated by the symptoms or abstain from following recommended and validated preventive measures.

As judiciously writes the editorialist: "In the critical evaluation of any screening test, one must answer the following questions: whether the test detects an early disease process, whether appropriate available intervention is most effective when applied early, whether testing improves outcomes in the screened population (as well as the number of tests required to find a preclinical disease) and whether patients are harmed by the screening test. (4).

"APPROPRIATE" USE OF CARDIOVASCULAR IMAGING

According to the study we have just discussed, we could say that TTE use is not convenient, adequate or appropriate in asymptomatic individuals. We therefore introduce ourselves in the "Appropriateness criteria", where in a series of documents, ad hoc groups define the usefulness of a cardiovascular procedure (in our case cardiovascular imaging) in relation with specific clinical questions, in order to define which, if any, procedure (imaging test) is indicated to help determine the diagnosis, treatment, or outcome.

The methodology adopted and recently published by the association of the American College of Radiology (ACR) and the American College of Cardiology Foundation (ACCF) accomplishes this aim through the application of systematic evidence reviews integrated with expert opinion by means of a rigorous Delphi process. These documents attempt to establish an evidence-based practical conduct that will provide treating physicians, imaging laboratories, interpreting physicians, patients, and those responsible for health care policies with a tool to make optimal use of cardiovascular imaging.

Relevant methods include echocardiography, radionuclide images, cardiac magnetic resonance, cardiac computed tomography and invasive coronary angiography. The optimal use of these procedures for specific clinical scenarios is unclear, and provides the core for the development of "appropriateness" recommendations. (5)

The reasons for the growth of imaging use are manifold; however, the improvement subjectively perceived by physicians, technicians and the patient is the main factor. But at the same time it has been questioned due to the wide geographical variability in the use of cardiovascular images, not explained by demographic difference or risk factors.

This methodology began in 1993 when the ACR developed "Appropriateness criteria" to assist physicians in the adequate or appropriate decision making of the images they requested. As of June 2012 there were 180 clinical scenarios of "Appropriateness Criteria", which are updated every 2 years.

The ACCF initiated the development of "Appropriateness criteria" in 2004 and as of 2010 six documents of appropriate use, each encompassing 50 to 200 different clinical scenarios, had been published.

The "Organizational structure" of an "Appropriateness criteria" document has four categories with very specific functions.

1. **The Oversight Committee** (12-16 members) conducts methodological oversight, defines the scope of the document and supports continuity of the whole process. The committee is also responsible for the selection of individuals who will be in the

three remaining panels: writing, review and rating, and for providing survey and approval at each stage of the process.

2. **The Writing Panel** (8-10 members) identifies clinical indications that are relevant for clinical decision scenarios and the use of cardiovascular imaging. It evaluates and categorizes the literature and constructs narratives and evidence tables for each indication or scenario.
3. **The Review Panel** (20-40 members) provides critical review and recommendations on the refining of the Writing Panel document before the rating process.
4. **The Rating Panel** (15-19 members) reviews the narratives and evidence tables and rates the appropriateness of imaging use for specific clinical indications. Each panel member performs an independent initial round of rating of each indication or scenario, followed by an in-person meeting to discuss and further refine the indications and evidences. If necessary, a second and even a third round of rating and appropriateness are made to end the rating criteria for the appropriate use of each clinical scenario.

   A score of 1 to 9 is used:

   1, 2 or 3 "rarely appropriate" (exceptions must have documentation of clinical reasons).

   4, 5 or 6 "maybe appropriate" (the procedure may be acceptable and may be reasonable for the indication).

   7, 8 or 9 "appropriate" (the procedure is generally acceptable and is generally reasonable for the indication).

   NOTE: Sufficient agreement is achieved for a topic if ≥ 60 % of the rating score classification falls within one of the three categories of "appropriate", "maybe appropriate "or "rarely appropriate"

APPROPRIATE USE AND CLINICAL IMPACT OF TRANSTHORACIC ECHOCARDIOGRAPHY

Let us now see how a guide of "Appropriate Use Criteria for Echocardiography", (6) developed in order to improve patient care and health outcomes, performs for the method that is responsible for almost half of all cardiac imaging studies.

Matulevicius et al. (7) compared the clinical impact of TTE with the echocardiography "Appropriateness criteria" rating of 2011 (6), since the association between TTE, "Appropriateness criteria" and their clinical impact had not been well explored.

The study is a retrospective review of medical records from 535 consecutive TTE (April 1 to 30, 2011) conducted in an academic medical center. Transthoracic echocardiographic studies were classified by two cardiologists "blinded" to clinical impact as "appropriate" (score 7-9), "uncertain" (score 4-6) or "inappropriate" (score 1-3), and were assessed for clinical impact by 2 cardiologists "blinded" to the "Appropriateness criteria" who retrospectively reviewed the electronic medical records and defined one of the three mutually exclusive categories: 1) active change in clinical care, 2) continuation of current care, or 3) no recommendation for care.

They found that 57% of TTE were from inpatients and orders had been mostly requested by internal medicine (38.5%) and cardiology (32.2%). Based on the 2011 criteria (6) 91.8% of TTE were appropriate, 4.3% were inappropriate and 3.9% were uncertain. The 10 most frequent "Appropriateness criteria" accounted for 66.5% of all TTE.

But despite the vast majority of echocardiograms were appropriate, only 1 in 3 TTE (31.8 %) resulted in an active change in clinical care, nearly 1 in 2 (46.9%) continued with the current care, and one in 5 (21.3%) had no change in care.

It is of great interest to mention that there is no significant association between appropriate and inappropriate TTE and active change in clinical care. These data suggest that the "Appropriateness Criteria" has not met the expected outcome: to have "a significant impact on physician decision making." (6)

The continuous growth of TTE since the publication of the "Appropriateness criteria "could neither be stopped nor the percentage of TTE classified as appropriate changed, prior to the publication in 2000 and after the publication of 2007 (87% vs. of 85%, p = 0.58) (8), indicating the absence of "Appropriateness criteria" use in clinical practice or the lack of sensitivity and specificity of the criteria developed.

Because the technical panels and the document writing group were mainly formed by physicians specialized in imaging, including many echocardiography experts, the consensus rating represents the current view of those focused on imaging and is unlikely to discourage the current practice by which TTE is ordered, resulting in a more liberal use of echocardiography.

As Armstrong and Eagle (9) argue, a retrospective review, without direct contact with the patient's physician, where an outsider looks for the "Appropriateness criteria", is a dichotomous decision that excludes the unknown shades and always involves a good faith effort to determine the indications for a study as perceived by the requiring physician. The same limitation appears in the evaluation of clinical impact, where it can be very difficult to define with certainty the absence of the clinical impact of the echocardiogram. For example, criterion 15 (evaluation of suspected pulmonary hypertension) resulted in active health care change in 10 of 22 people; nevertheless, the result in the other 12 of 22 people where no changes were made is also important if it allowed rejecting a well-based suspicion of primary or secondary pulmonary hypertension due to the presence of a disease that could produce it.

For future studies seeking the effectiveness of "Appropriateness criteria", they should at least be performed prospectively, looking for the impact of echo-

cardiography and how it affects medical decisions.

TO RECONSIDER CLINICAL CARDIAC IMAGING USEFULNESS

Instead of Appropriateness criteria, why not randomize the strategy?

We are under the burden and overwhelmed by a plethora of new expensive diagnostic tests, and all interested parties are struggling to try to define the criteria by which the use of these tests is suitable or "appropriate". As we have seen "Appropriateness criteria" are in the end decided on the basis of expert opinion and some circumstantial evidence concerning the performance characteristics of diagnostic tests (sensitivity, specificity, accuracy). "However, as proposed by Ioannidis (10) - what really matters in the end, refers to what happens to patients undergoing tests. What other tests, invasive procedures, or treatments were ordered or aborted based on the results? How were the major clinical events, quality of life, or even survival (for serious conditions) affected? Often data to answer these questions are scarce or nonexistent."

To restrict TTE performance to the "appropriate" indications, 10 highly prestigious scientific societies and the best experts met in an extraordinary effort, with full transparency in methods, processes and conclusions, to prepare a meticulous and extensive list of 202 TTE indications, 97 of which were classified as "appropriate" , 34 as "uncertain" and 71 as "inappropriate" .

This effort to ensure a minimum of inappropriate TTE and thus optimize the results is not reflected in the recent work of Matulevicius (7) we have just discussed. Although 91.8 % of the echocardiograms were undoubtedly appropriate according to these criteria, and consistent with similar previous studies, it only shows that the appropriate use guideline only reflects everyday practice, as it was mostly based on the opinion of echocardiography experts. Furthermore, active changes are not significantly different in appropriate or inappropriate TTE, so the Matulevicius study showed that the concepts of "appropriate use" and "usefulness" may diverge considerably.

In fact, our uncertainty about the clinical benefits of TTE and other cardiac imaging methods is still very high, and active changes in patient management after a TTE is a substitute of little use, as an active change may lead to the emergence of new complications by choosing perhaps unnecessary diagnostic or therapeutic interventions, as in the screening of asymptomatic subjects we mentioned at the beginning of this letter. Sometimes an echocardiogram that simply reassures the patient and the doctor to continue the same course may be harmful, if the TTE showing, for example, normal left ventricular function leads to rule out a clear dyspnea that may indicate heart failure with preserved function.

In the absence of comparative data in ideal randomized TTE clinical trials, we cannot say whether the changes in the action plan are equated with clinical usefulness. If this inference is true, the simple creation of a list of appropriate indications does not mean that their use will lead to some benefit for the patient. Is it not possible that the justification of "appropriate" in American medicine will result in a defensive medicine , where the documentation of "appropriate" has the meaning of "refundable" without really helping the patient ?

We should agree that the greatest difficulty in declaring an indication of TTE as appropriate, simply lies in our lack of good evidence of how and when to use the echocardiogram, as in the appropriate use guidelines divided in only two scenarios the level of evidence is "A" (implantation of an ICD with ejection fraction < 35 % and in bacterial endocarditis) and the vast majority have little or no evidence.

CONCLUSIONS

As Ioannidis (10) states in the title: "randomized clinical trials of diagnostic tests should be encouraged. These trials need to focus on specific clinical scenarios, and experts on the area may be consulted on how to define them."

The number of participants in clinical trials for commonly used procedures (echocardiography) and for commonly proposed indications should not be limited, and should follow the example of the study on the value of echocardiography in asymptomatic individuals where 6861 persons were included in a single evaluation.

The problem faced by the clinician is not only whether or not a test is needed, but whether a test (new) is better than another (old), or if a diagnostic sequence strategy (in series or in parallel) is preferable to another. Therefore, randomized clinical trials should be designed to answer these questions.

Until these definitive studies are performed, the distinction between "appropriate" versus "clinically useful" will remain very difficult.

Dr. Hernán C. Doval[MTSAC]
Director of the Argentine Journal of Cardiology

REFERENCES

1. Greenland P, Alpert JS, Beller GA, Benjamin EJ, Budoff MJ, Fayad ZA, et al. American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. Circulation 2010;122:e584-e636. http://doi.org/fk98sn

2. Andrus BW, Welch HG. Medicare services provided by cardiologists in the United States: 1999-2008. CircCardiovascQual Outcomes 2012;5:31-6. http://doi.org/fzdrk8

3. Lindekleiv H, Løchen M-J, Mathiesen EB, Njølstad I, Wilsgaard T, Schirmer H. Echocardiographic screening of the general population and long-term survival. A randomized clinical study. JAMA Intern Med 2013;173:1592-7. http://doi.org/rc2

4. Michos ED, Abraham TP. Echoing the appropriate use criteria. The role of echocardiography for cardiovascular risk assessment of the asymptomatic individual. JAMA Intern Med 2013;173:1598-9. http://doi.org/rc3
5. Carr JJ, Hendel RC, White RD, Patel MR, Wolk MJ, Bettmann MA, et al. 2013 appropriate utilization of cardiovascular imaging: a methodology for the development of joint criteria for the appropriate utilization of cardiovascular imaging by the American College of Cardiology Foundation and American College of Radiology. J Am Coll Cardiol 2013;61:2199-206.
6. 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; Society for Cardiovascular Magnetic Resonance, Douglas PS, Garcia MJ, Haines DE, Lai WW, Manning WJ, Patel AR, Picard MH, et al. ACCF/ASE/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. doi: 10.1016/j.jacc.2010.11.002
7. Matulevicius SA, Rohatgi A, Das SR, Price AL, Deluna A, Reimold SC. Appropriate use and clinical impact of TTE. JAMA Inter Med 2013;173:1600-7. http://doi.org/rc4
8. Rahimi AR, York M, Gheewala M, Markson L, Hauser TH, Manning WJ. Trends in outpatients TTE: impact of appropriateness criteria publication. Am J Med 2011; 124:740-6. http://doi.org/crjtqg
9. Armstrong W, Eagle KA. Appropriate use criteria in echocardiography. Is no change the same as no benefit? JAMA Intern Med 2013; 173:1609-10. http://doi.org/rc5
10. Ioannidis JPA. Appropriate vs clinically useful diagnostic tests. JAMA Intern Med 2013; 173:1607-9. http://doi.org/rc6