Report on a national echocardiography quality-control exercise

Kevin F. Fox1,3*, Andrew Porter2, Beth Unsworth2, Tim Collier3, Graham Leech1, and Jamil Mayet1,3

1British Society of Echocardiography, DBC, 10–16 Tiller Road, Docklands, London E14 8PX, UK; 2Imperial College Healthcare NHS Trust, UK; and 3Imperial College, London, UK

Aims To assess the feasibility of conducting a large-scale quality-control exercise on the echocardiographic grading of mitral regurgitation (MR).

Methods and results One hundred and fifty-two practising cardiologist and sonographer echocardiographers attended a dedicated session within a national meeting and were asked to review echocardiographic images of five case studies and evaluate the severity of MR using a six-point scale. The group's overall evaluation was analysed together with the variation in grading of severity. The proportion of gradings of MR more than one grade either side of the mode was <10% in all but one case, and <10% were inaccurately evaluated as inside categories of severe when the modal grading was outside and vice versa. However in a case where a single grading difference had important clinical implications, substantial variability was seen.

Conclusions Conducting a large-scale quality-control exercise in the assessment of mitral regurgitation is feasible. Overall results suggesting reasonable consistency in reporting may hide substantial clinically relevant variability. It is essential that increasing importance is attached to the development, conduct, and analysis of quality control within echocardiography if it is to maintain and extend its role as a key investigation for patients with heart disease.

KEYWORDS
Echocardiography; Mitral regurgitation; Quality Control

Introduction
The delivery of high-quality echocardiography requires skilled individuals working in high-quality echo departments with a continuous process of quality assurance. The importance of quality assurance (also called quality control) is increasingly recognized but rather under developed. Criteria for high-quality echo departments have been published, which include the requirement for quality-control programmes, and systems of accreditation for individuals and departments have been established (www.bsecho.org).1 While research studies may set criteria for the quality of performance of echo studies, and may monitor participating centres, few routine clinical parameters are published. An exception is the British Society of Echocardiography statement on the accuracy and reproducibility of echocardiographic studies performed for patients receiving Trastuzumab chemotherapy.2 This identifies a standard for reproducibility of less than or equal to 10% variability of ejection fraction measurements for departments performing such echocardiograms. No such parameters exist for the assessment of valvular regurgitation and no published data exist on the current variability in the reporting of valvular regurgitation.

The British Society of Echocardiography (BSE) has specifically, during the last 2 years, promoted the expansion of quality control within UK echocardiography departments. As a unique experiment a large-scale, quality-control exercise was conducted during the BSE Annual Scientific meeting in October 2007.

Methods
The quality-control exercise was designed to assess the variability in the grading of mitral regurgitation.

Five consecutive studies, with any degree of mitral regurgitation, performed on a Phillips IE 33 machine at Charing Cross Hospital in West London were selected.

2D, M Mode, colour, pulse wave, and continuous wave images were acquired following the standard methods in use in the department for routine clinical studies. This included quantification using PISA but not 3D imaging. The sonographers were not aware the
The results from the initial eight sonographers and clinicians are shown in Table 1.

Table 1: Single hospital sonographer and cardiologist gradings of mitral regurgitation in five cases

<table>
<thead>
<tr>
<th>Operator</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>II</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>III</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>IV</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>V</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

The initial exercise was completed by eight sonographers and physicians at Charing Cross Hospital.

The required data and grading are as follows:

- Nil or Trivial
- Mild
- Mild-to-Moderate
- Moderate
- Moderate-to-Severe

Results

One hundred and fifty two people participated in the complete exercise. The breakdown of attendees was 75% sonographers, 25% physicians, and 55% fully accredited echocardiographers, 45% not accredited (including in training).

The initial exercise was completed by eight sonographers and physicians at Charing Cross Hospital.

All gradings adhere to the following code:

- Nil or Trivial
- Mild
- Mild-to-Moderate
- Moderate
- Moderate-to-Severe
- Severe

The results from the initial eight sonographers and physicians are shown in Table 1.

Table 2: Details of cases and results of gradings by 152 people for the five cases

<table>
<thead>
<tr>
<th>Grade</th>
<th>Case 1 (%)</th>
<th>Case 2 (%)</th>
<th>Case 2extra (%)</th>
<th>Case 3 (%)</th>
<th>Case 4 (%)</th>
<th>Case 5 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nil or trivial</td>
<td>1 (0)</td>
<td>2 (1)</td>
<td>42 (31)</td>
<td>69 (46)</td>
<td>42 (28)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Mild</td>
<td>2 (1)</td>
<td>79 (54)</td>
<td>93 (69)</td>
<td>72 (50)</td>
<td>92 (62)</td>
<td>32 (21)</td>
</tr>
<tr>
<td>Mild-to-moderate</td>
<td>3 (1)</td>
<td>43 (29)</td>
<td>0 (0)</td>
<td>5 (3)</td>
<td>13 (9)</td>
<td>66 (43)</td>
</tr>
<tr>
<td>Moderate</td>
<td>4 (1)</td>
<td>11 (7)</td>
<td>0 (0)</td>
<td>2 (1)</td>
<td>2 (1)</td>
<td>42 (28)</td>
</tr>
<tr>
<td>Moderate-to-severe</td>
<td>5 (5)</td>
<td>61 (41)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>9 (6)</td>
</tr>
<tr>
<td>Severe</td>
<td>6 (6)</td>
<td>77 (52)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (1)</td>
</tr>
</tbody>
</table>

Views shown:

- PSLA, PSLAC, A4C, A2C, A4C CW, PISA
- PSLA, PSLAC, PSSA, A4C, A4C C, ASC CW, A4C CW
- PSLA, PSLAC, PSSA, A4C, A4C PW, ASC PW, A4C CW
- PSLA, PSLA C, PSSA, A4C, A4C PW CW, A4C CW, PISA
- PSLA, PSLA C, PSSA, A4C, A4C PW, A4C PW, A4C CW, PISA

Final Diagnosis/Plan:

- Dilated cardiomyopathy with secondary moderate-to-severe MR
- Mild-to-moderate MR
- MV surgery
- Mild MR medical therapy
- Mild MR medical therapy
- Dilated cardiomyopathy with secondary moderate MR

PS, parasternal; LA, long axis; SA, short axis; A4C, apical 4 chamber; A2C, apical 2 chamber; ASC, apical 5 chamber; C, Colour Doppler; CW, Continuous Wave Doppler; PW, pulse wave; PISA, proximal isovolumic surface area data; MR, mitral regurgitation.
Table 2 shows the gradings of the 152 people and the final diagnosis and plan made by the multidisciplinary clinical team managing the patient.

The results from the national quality-control exercise are shown in Figures 1–6, additional results are shown in Table 3.

There was variability in the grading of mitral regurgitation. However in four out of five cases <10% of gradings differed from the modal (most common) grading by more than 1 grade. In Case 2 (rheumatic mixed mitral valve disease) with a yes/no answer, ‘is the mitral regurgitation too severe for balloon valvuloplasty?’, the grading was split 31% no, 69% yes. Case 5 had the greatest spread partly because the mitral regurgitation was in the centre of the possible range (moderate). The most common grading by the initial eight sonographers was the same or at most one grade different from the results of the larger group of 152.

Discussion

Accurate reporting of echocardiograms, like all investigations, is mandatory for best care. To achieve this, one needs appropriate and consistently trained individuals practicing within a high-quality environment. However even this will not ensure effective practice without adequate quality assurance.

Quality assurance or quality-control programmes are embraced within industry and in many areas of healthcare.
But in clinical practice, and certainly echocardiography, introduction has been patchy. The literature on such programmes is equally sparse. Hidden within many trials may be information on within and between centre reproducibility of echocardiographic measures, perhaps with additional core lab adjudication, but there is a lack of studies on quality-control programmes in clinical echocardiography with no published studies on variability in reporting of valve lesions.

The first aim of this project was to establish if it was possible to conduct a large-scale quality-control exercise. This was achieved. A relatively straightforward, within department exercise was reproduced in a large forum.

The second aim was to evaluate the accuracy and consistency of a large group of practicing echocardiographers in reporting one of the common parameters, the severity of mitral regurgitation. While quantitative parameters are available to evaluate mitral regurgitation severity, it is still widely assessed by the amalgamation of a variety of quantitative and qualitative measures synthesized by the echocardiographer to produce a final grading.

The next major hurdle is determining what is an acceptable spread of grading of mitral regurgitation in the absence of any published criteria. We do not propose a 'correct' answer for the severity of mitral regurgitation in each case. If such an answer were necessary, the use of panels of experts would probably be an appropriate methodology. Given the subjective elements to evaluation some spread is inevitable. The difficulty is that the significance of variability in grading fluctuates. As a core measure, we have used the proportion of people who are greater than one grade from the most commonly assigned grade or mode grade (so if a case is felt overall to show moderate regurgitation, this is the proportion who did not grade it as mild-to-moderate, moderate, or moderate-to-severe). In all but one case this proportion was <10% which is seemingly reassuring. However these simple parameters hide some clinically very important variability. The second case, one of rheumatic mitral stenosis with mitral regurgitation illustrates this well since the severity of mitral regurgitation is a key parameter in suitability for balloon valvuloplasty. It is therefore of concern that the echocardiographers showed substantial variability in this assessment.

A further significant concern is that the group data may hide individuals consistently under or over grading. Many patients are followed longitudinally for their valvular heart disease. Such monitoring is severely compromised if grading is inconsistent. A limitation of the study is that we were unable to track individual or groups of individuals in the same echo laboratory’s grading patterns, or to stratify by experience and qualifications of the participants.

It is important to note that this exercise differs from normal clinical practice in many ways. We showed pre-recorded clips and stills. While they were chosen to provide the full range of parameters used to assess mitral regurgitation, echocardiographers may be more consistent and accurate in assessing mitral regurgitation on studies they had themselves performed. In addition, most echo laboratories have systems for joint review of critical cases, and this is intended to improve consistency and accuracy. So the mitral stenosis case would usually be reviewed by a team of surgeons, imaging cardiologists, and interventional cardiologists to determine the mode of treatment.

More generally the decision on whether a valve needs operative correction is not made solely on the basis of the echo. It is made following a synthesis of clinical data, echo, and other investigations assessed by a team of physicians and surgeons. However the reporting of ‘severe’ (i.e. a grading of ‘moderate to severe’ or ‘severe’) stimulates thoughts of surgical intervention. Encouragingly in Case 1 only 8% did not mention ‘severe’ and in Case 5 only 7% did mention ‘severe’ in their gradings, reporting that might have inappropriately diverted a patient away from, or towards a surgical treatment plan, respectively. However laboratories should not rely on such checks and balances without data from quality-control exercises such as this one.

**Conclusion**

A large-scale, quality-control exercise to evaluate variability in the grading of mitral regurgitation is feasible. While the results can be interpreted as <10% of echocardiographers grade mitral regurgitation significantly differently from the modal grading of a large group, within this may be substantial variability with important clinical
consequences. It is essential that increasing importance is attached to the development, conduct, and analysis of quality control within echocardiography, if it is to maintain and extend its role as a key investigation for patients with heart disease.

Conflict of interest: none declared.

References


2. Fox KF. The evaluation of left ventricular function for patients being considered for or receiving Trastuzumab (Herceptin) therapy. Br J Cancer 2006;95:1454.