Left ventricular ejection fraction (LVEF) is a clinically accepted measurement for assessment of left ventricular systolic function. The correct measurement of LVEF is clinically important for reliably estimating the severity and prognosis of many cardiovascular diseases (1, 2). It is also used to determine the optimal treatment strategy, including the indication of an implantable cardioverter defibrillator or biventricular pacing (3). LVEF can be measured quantitatively by left ventriculography, M-mode, two-dimensional (2D) and three-dimensional (3D) echocardiography, radionuclide imaging techniques, multislice computerized tomography (MSCT), and cardiac magnetic resonance imaging (CMRI) as well as estimated by visual assessment on echocardiography.

LVEF measurement is accurate and highly reproducible with CMRI. CMRI with its non-invasive nature and lack of radiation risk outperform other techniques in terms of measurement of left ventricular volume and ejection fraction owing to its better soft tissue contrast and high spatial resolution, ability to assess all the chambers, and the ability to visualize the entire cardiac cycle. CMRI is particularly useful in patients with poor image quality on echocardiography.
myocardial segments, and independence from geometric assumptions. All these factors make CMRI the gold standard for the measurement of LVEF (4); however, CMRI is not always easily applicable. Echocardiography is a faster and more accessible examination and therefore is used more frequently.

Modern echocardiographic recommendations advise researchers to follow the Simpson method in calculating LVEF (5). However, visual evaluation can be considered as an important alternative technique in estimating LVEF by echocardiography with a rapid assessment of left ventricular systolic function in clinical practice, especially in laboratories with a high workload, except for critical decisions that depend on quantitative LVEF measurement (6). Several studies, with some limitations, compared visually estimated LVEF (ve-LVEF) and quantitatively measured LVEF (qm-LVEF) measured with left ventriculography, radionuclide techniques; or, in a few studies, CMRI. It will be the most appropriate approach for each laboratory estimating LVEF visually to demonstrate whether these assessments agree with quantitative methods.

The main objective of our study was to determine the accuracy of visual analysis of LVEF in comparison with the quantitative gold standard CMRI using contemporary statistical methods. Another aim was to evaluate the intraobserver and interobserver agreement associated with ve-LVEF.

METHODS

Subjects and study design
A total of 54 patients (17 men and 37 women), aged between 22 and 73 years, undergoing CMRI examinations because of various clinical preliminary diagnoses were enrolled in our study.

CMRI was performed on all our patients at the radiology department of our center. The patients were then sent to our echocardiography laboratory to record transthoracic echocardiography images on the same day. There was no change in the clinical findings and treatment protocols of patients during this time period.

In this study, considering ease of use in clinical practice, visually estimated values of LVEF in echocardiography were planned to be compared with other quantitative methods. We planned to compare ve-LVEF with echocardiographic Simpson method and the gold standard technique CMRI as reference methods. However, despite video records of echocardiographic images that had acceptable quality, endocardial borders were not that clear in the still frames in some patients. Therefore, we compared visual estimations only with CMRI, the gold standard method.

The study was approved by the Local Ethics Committee in accordance with the Declaration of Helsinki and good clinical practice. Informed consent was obtained from all the patients.

Echocardiographic evaluation
Echocardiographic examinations of the patients were performed using a 3.5 MHz cardiac probe by Vivid 55 (GE Medical Systems, Milwaukee, MI, USA). Parasternal long and short axis images of the patients (at basal, middle, and apical levels), apical two and four chamber images, and apical long axis images were recorded. Echocardiographers were randomly selected. The evaluation of the ejection fraction was performed only visually by two independent echocardiographers without using any quantitative method. During this evaluation, the image quality was classified as good, moderate, and poor. After at least 1 week from the initial assessment, one of the observers visually re-estimated the LVEF value from these images. The first and second estimations made by the same echocardiographer were used to assess the intraobserver variability.

CMR image acquisition and image analyses
Cardiac MRI scans were performed with a 1.5 Tesla MR device (Optima MR 450 W, GE) with the cardiac coil at magnetic resonance imaging unit of the radiology department.

Axial, coronal, and sagittal planar triggered, breathable, “single shot” pilot images were obtained. The axial plane FIESTA (b-SSFP) images of the heart anatomy, including the area from upper mediastinum to the liver dome, were obtained for a demonstration of non-cardiac pathologies. Real-time pilot imaging revealed vertical long axis, horizontal long axis, short axis, and four chamber images. Vertical long axis FIESTA images of the left ventricle were provided by real-time images. Horizontal long-axis FIESTA images were obtained by correlation of vertical long axis images with real time short axis pilot images. Short-axis FIESTA images with 1 cm intervals (6–8 mm slice with a 2–4 mm gap) were obtained by vertical, horizontal long axis, and real time four-chamber pilot images. The images were taken using ECG triggered method with the patient holding breathing in expiration. When each FIESTA image was obtained, a manual shim was made corresponding to half of the field-of-view (FOV) to prevent flow artifacts.

Left ventricular volume and function were measured with a specially equipped computer program (GE, AW-Report card 4.0) blinded to the echocardiographic results. For calculation of LVEF values, end-diastolic/systolic endocardial margins were visualized semi-automatically in all sections from apex to baseline in short axis images, and then corrections were made manually at all phases. End-diastolic and systolic volumes and the values of ejection fractions were measured using this method.
Statistical analysis
Sample size calculation was first based on an estimation of intraclass correlation (ICC) coefficients of $p_0=0.6$ and $p_1=0.8$. Taking the alpha error as 0.05 and beta error 0.20, the required sample size was calculated as 39 patients. The sample size was calculated on the basis of the preliminary data obtained during this study period, in which ICC coefficient values were $p_0=0.8$ and $p_1=0.9$; and the required sample size was at least 46 persons. We enrolled 54 persons in the study.

ICC values and Bland–Altman plots were used to assess interobserver and intraobserver agreement. The ICC value ranges from 0 to 1, and it is accepted that agreement is excellent for the values of 0.95–1, high for 0.85–0.94, and moderate for 0.70–0.84 (7).

Two categorizations were applied to LVEF values. The three-group categorization included the LVEF categories of ≤35%, 36%–54%, and ≥55%. To assess normal or reduced LVEF categories, two-group categorization was constituted as LVEF values of <55% and ≥55%. The agreement between the categorized LVEF groups was assessed with unweighted and linearly weighted kappa statistics. The kappa values were interpreted as follows: <0.20 poor or slight; 0.21–0.40 weak; 0.41–0.60 moderate; 0.61–0.80 good; 0.81–1 almost perfect agreement.

RESULTS
The characteristics of the patients are given in Table 1. The mean age of the study group was 47.7±15.7 years, and 37 (68.5%) of them were women. The LVEF values measured by CMRI were between 21.4% and 80.3% with an average of 56.4±14.5%.

Agreement between visually assessed and CMRI calculated LVEF
There was a good agreement between ve-LVEF and CMRI calculated LVEF values [ICC: 0.93 (95% CI 0.88–0.96)]. The Bland–Altman plot showed that visually estimated LVEF was on average 0.6 percentage points (95% CI −10.5 and +9.3) lower than the values calculated from CMRI (Fig. 1). The agreement was also related to the image quality; the limits of agreement were slightly better when the image quality is good (Fig. 1). It was also noted that the difference between the two methods was distributed in a slightly narrower range in cases with the LVEF values of less than 55%–60% compared with those of higher values.

Table 1. Mean age of the patients, sex distribution, and LVEF values

| Age (mean ± SD) (years) | 47.7±15.7 |
|-------------------------|-----------|
| Sex F/M [n (%)]         | 37 (68.5%)/17 (31.5) |
| LVEF (mean ± SD)        | 56.4%±14.5% |
| LVEF category 1         |           |
| LVEF ≤ 35% n (%)        | 5 (9.3)   |
| LVEF 36%–54% n (%)      | 16 (29.6) |
| LVEF ≥ 55% n (%)        | 33 (61.1) |
| LVEF category 2         |           |
| LVEF < 55% n (%)        | 21 (38.9) |
| LVEF ≥ 55% n (%)        | 33 (61.1) |
| LVEF - left ventricular ejection fraction | |

Agreement between the two methods was evaluated using the kappa statistics for the subgroups formed by categorizing the LVEF values. For three-category groups (LVEF categories of ≤35%, 36%–54%, and ≥55%), unweighted kappa statistics was 0.71 (standard error: 0.09), and linearly weighted kappa statistics was 0.76 (standard error: 0.07) (Table 2), suggesting that the agreement between the two methods was good. When the kappa statistics was assessed for the two-category groups (LVEF categories of <55% and ≥55%), the kappa value increased to 0.80, suggesting a perfect agreement between the two methods (Table 3). Visual estimation gave discrepant results in only 1 patient among those with CMRI-calculated LVEF ≥55% (agreement 32/33; 97.0%).

Table 2. Kappa statistic evaluating the agreement between the two methods in three-category groups

| CMRI LVEF | ≤35% (n) | 36%–54% (n) | ≥55% (n) | n (%) |
|-----------|----------|-------------|---------|-------|
| Visual LVEF |          |            |         |       |
| ≤35%       | 3        | 2           | 0       | 5 (9.3)|
| 36%–54%    | 2        | 11          | 1       | 14 (25.9)|
| ≥55%       | 0        | 3           | 32      | 35 (6.8)|
| n (%)      | 5 (9.3)  | 16 (29.6)   | 33 (61.1)| 54    |
| Kappa      | 0.71 (standard deviation: 0.09) | Linearly weighted kappa: 0.76 (standard deviation: 0.07) |

Table 3. Kappa statistic evaluating the agreement between the two methods in two-category groups

| CMRI LVEF | <55% (n) | ≥55% (n) | n (%) |
|-----------|----------|---------|-------|
| Visual LVEF |        |         |       |
| <55%      | 18       | 1       | 19 (35.2)|
| ≥55%      | 3        | 32      | 35 (64.8)|
| n (%)     | 21 (38.9)| 33 (61.1)| 54    |
| Kappa     | 0.80 (standard deviation: 0.08) |

CMRI - cardiac magnetic resonance imaging; LVEF - left ventricular ejection fraction
and in 3 patients among those with CMRI-calculated LVEF <55% (agreement 18/21; 85.7%) (Table 3). However, when the calculations were made for ve-LVEF, 18 (94.7%) of 19 patients with estimated LVEF <55%, and 32 (91.4%) of 35 patients with estimated LVEF ≥55% were correctly classified.

Evaluation of intraobserver agreement

Intraobserver agreement between the visual estimations were excellent [ICC: 0.96 (95% confidence interval 0.94–0.98)]. In the Bland-Altman plots, the mean difference between the 2 visually estimated LVEF values was +0.8 percentage points (95% CI −5.3 to +7.8). For patients with good image quality, the limits of agreement were better (Fig. 2).

Intraobserver agreement was also assessed for ve-LVEF categories. For the three-category groups (LVEF categories of ≤35%, 36%–54%, and ≥55%), unweighted kappa statistics was 0.85 (standard error: 0.06), and linearly weighted kappa statistics was 0.88 (standard error: 0.05), suggesting that intraobserver agreement between ve-LVEF categories was very good. When the kappa statistics was calculated for the two-category groups (LVEF categories of <55% and ≥55%), the value increased to 1.00, suggesting a perfect intraobserver agreement between ve-LVEF categories.

Evaluation of interobserver agreement

Interobserver agreement between the visual estimations were high [ICC: 0.91 (95% confidence interval 0.80–0.95)]. In the Bland-Altman plots, mean difference between the two ve-LVEF values was −2.8 percentage points (95% CI −12.9 to +7.3). For patients with good image quality, the limits of agreement were slightly better (Fig. 3). For three-category groups (LVEF categories of ≤35%, 36%–54%, and ≥55%), unweighted kappa statistics for interobserver agreement on ve-LVEF categories was 0.89 (standard error: 0.06), and linearly weighted kappa statistics was 0.91 (standard error: 0.05), suggesting a very good interobserver agreement. When the kappa statistics was calculated for the two-category groups (LVEF categories of <55% and ≥55%), the value increased to 0.92, suggesting a perfect interobserver agreement.

DISCUSSION

This study shows a good agreement between the ve-LVEF on echocardiography and qm-LVEF by CMRI. The intra- and inter-observer agreements between the ve-LVEF were also good, especially for patients with a good image quality and LVEF values of approximately <55%–60%.

Given the important clinical consequences associated with an accurate assessment of LVEF, quantitative assessment of LVEF should remain the clinical standard for clinicians. However, in many patients, exact LVEF values are not so critical, and approximate values with acceptable accuracy is adequate. Therefore, ve-LVEF seems to be an effective alternative approach, especially for busy echocardiography laboratories as it is easy and time efficient. However, it is crucial that the reliability of this approach be determined not only for general purpose but also for each laboratory that may want to use this method. The objective of our study was to assess the agreement between ve-LVEF and qm-LVEF with CMRI.

Several studies have compared ve-LVEF with qm-LVEF with left ventriculography, radionuclide angiography, echocardiography using automated biplane EF method, and CMRI (8–12). These studies showed an acceptable correlation between visual estimation and quantitative LVEF, and the correlation was affected by image quality. In most of these trials, Pearson or Spearman correlation coefficients were used to assess the agreement between the measurements. However, these coefficients alone are not appropriate measures for assessing the agreement and might be misleading (7). The strength of our study was assessing the agreement using appropriate statistical methods such as ICC, Bland-Altman plots, and kappa statistics.

In two different comparative studies performed with CMRI, it was noted that although the values of correlation coeffi-
cient were high in both studies, the agreement between the two methods was not very good and that this was owing to underestimation of LVEF with visual assessment compared with that with CMRI (11, 12). To determine precisely about how the margin of error changes in different LVEF values, we also used Bland–Altman plots in our trial, just as in these two studies.

The Bland–Altman plot showed that ve-LVEF was, on average, 0.6 percentage points lower than the values calculated with CMRI. As expected, the LVEF difference varies according to the image quality. In case of good image quality, the difference between CMRI and visual assessment is distributed in a narrower range, especially in those with LVEF of less than 60%. In the echocardiographic examinations performed as part of routine clinical practice, the amount of error in LVEFs above 60% can be considered relatively less important, except in particular cases in which LVEF measurement should be performed more precisely, such as presence of indication of valve surgery or follow-up of patients receiving chemotherapy. In a systematic review comparing the reliability of echocardiographic parameters with other quantitative parameters in evaluating LVEF, confidence interval of the difference of LVEF evaluation between the two methods varied in the range ±19% to ±24% in studies involving mixed patient populations; although it varied between ±16% and ±18% in studies involving only patients with myocardial infarction (13). When we compared these data with those from our study, we observed that the confidence interval that reflects the LVEF difference was narrower in our study (±99%). The higher resolution and good image quality of recent sonography devices might play a role of decreasing the variability between the methods.

LVEF value is preferred as a continuous variable, especially when follow-up assessment is required; however, in many patients, the medical decision depends on whether LVEF is normal or low. Therefore, we analyzed the agreement after categorizing the LVEF values using clinically meaningful cut-offs. Kappa statistics demonstrated a good agreement between the two methods both in the three-category (LVEF categories of ≤35%, 36%–54%, and ≥55%) and in the two-category classifications (LVEF categories of <55% and ≥55%). Discrepancy was observed in only 4 patients when the cut-off value of LVEF was accepted as 55%; however, the agreement between the two methods was found to be very good when the cut-off value for the LVEF was taken as 60%. When the results were assessed from a sensitivity and specificity perspective (using qm-LVEF as the denominator), the correct classification was observed in 85.7% of patients with CMRI calculated LVEF <55% and in 97.0% of patients with CMR calculated LVEF >55%. When the results were assessed from a predictive values perspective (using ve-LVEF as the denominator), the correct classification was 94.7% and 91.4% among patients with “estimated” LVEF <55% and ≥55%, respectively. Although the low number of patients in each category is an important limitation, these results suggest an acceptable agreement both in low and normal LVEF.

In the visual evaluation of LVEF, interobserver and intraobserver agreements were also considered as an important factor in determining the reliability of the results. Blondheim et al. (14) reported that interobserver and interobserver agreements of ve-LVEF were good (ICC 0.72 and 0.78, respectively). We found that ICC value was 0.91 for intraobserver agreement and 0.96 for interobserver agreement. Kappa statistics was also very good for agreement between the two methods for LVEF categories (≥0.85).

**Study limitations**

There were several limitations in our study. As the number of patients with LVEF ≤35% was low in our study, the results in this group should be assessed with some caution. Second, these results may not be applicable to other centers and to echocardiographers who have less experience. Of note, there are studies in the literature indicating the reliability of visual assessment of LVEF depends on the experience and duration of the researcher's work (15).

**CONCLUSION**

Qm-LVEF should not be replaced with visual estimation when it is used in a critical decision. However, the visual approach for LVEF assessment may be used for rapid assessment of left ventricular function in clinical practice, particularly in patients with good image quality.

**Conflict of interest:** None declared.

**Peer-review:** Internally peer-reviewed.

**Author contributions:** Concept – K.E., M.K.; Design – K.E., M.K., İ.D.; Supervision – M.K., C.T.K., İ.D.; Fundings – E.P., V.K.; Materials – E.P., Ç.K.; Data collection &/or processing – M.K., D.M.G.U.; Analysis &/or interpretation – V.K., Ç.K., D.M.G.U.; Literature search – K.E., V.K., Ç.K.; Writing – K.E., M.K.; Critical review – M.K., C.T.K., İ.D.

**REFERENCES**

1. Writing Committee Members, Otto CM, Nishimura RA, Bonow RO, Carabello BA, Erwin JP 3rd, et al. 2020 ACC/AHA Guideline for the Management of Patients With Valvular Heart Disease: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. J Am Coll Cardiol 2021; 77: e25-197.

2. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Colvin MM, et al. 2017 ACC/AHA/HFSA Focused Update of the 2013 ACCF/AHA Guideline for the Management of Heart Failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. J Am Coll Cardiol 2017; 70: 776-803. [Crossref]

3. Ojo A, Tarig S, Harikrishnan P, Iwai S, Jacobson JT. Cardiac Re-synchronization Therapy for Heart Failure. Interv Cardiol Clin 2017; 6: 417-26. [Crossref]

4. Pontone G, Guaricci AI, Andreini D, Solbiati A, Guglielmo M, Mustaq S, et al. Prognostic Benefit of Cardiac Magnetic Resonance Over Transthoracic Echocardiography for the Assessment of Ischemic and Nonischemic Dilated Cardiomyopathy Patients Referred for the Evaluation of Primary Prevention Implantable Cardioverter-Defibrillator Therapy. Circ Cardiovasc Imaging 2016; 9: e004956. [Crossref]

5. Galderisi M, Cosyns B, Edvardsen T, Cardim N, Delgado V, Di Salvo G, et al.; 2016–2018 EACVI Scientific Documents Com-
mittee; 2016–2018 EACVI Scientific Documents Committee. Standardization of adult transthoracic echocardiography reporting in agreement with recent chamber quantification, diastolic function, and heart valve disease recommendations: an expert consensus document of the European Association of Cardiovascular Imaging. Eur Heart J Cardiovasc Imaging 2017; 18: 1301-10. [Crossref]

6. Liu D, Peck I, Dangi S, Schwarz KQ, Linte CA. Left Ventricular Ejection Fraction Assessment: Unraveling the Bias between Area- and Volume-based Estimates. Proc SPIE Int Soc Opt Eng 2019; 10955: 109550T. [Crossref]

7. Alpar R. Geçerlik ve Güvenirlik. In: Alpar R, editor. Spor, Sağlık ve Eğitim Bilimlerinden Örneklerle Uygulamalı İstatistik ve Geçerlik-Güvenirlik SPSS de Çözümleme Adımları ile Birlikte. Ankara: Detay Yayıncılık; 2020. p.439-73.

8. Mueller X, Stauffer JC, Jaussi A, Goy JJ, Kappenberger L. Subjective visual echocardiographic estimate of left ventricular ejection fraction as an alternative to conventional echocardiographic methods: comparison with contrast angiography. Clin Cardiol 1991; 14: 898-902. [Crossref]

9. Amico AF, Lichtenberg GS, Reisner SA, Stone CK, Schwartz RG, Meltzer RS. Superiority of visual versus computerized echocardiographic estimation of radionuclide left ventricular ejection fraction. Am Heart J 1989; 118: 1259-65. [Crossref]

10. Abazid RM, Abohamr SI, Smettei OA, Qasem MS, Suresh AR, Al-Harbi MF, et al. Visual versus fully automated assessment of left ventricular ejection fraction. Avicenna J Med 2018; 8: 41-5. [Crossref]

11. Sievers B, Kirchberg S, Franken U, Puthenveettil BJ, Bakan A, Trappe HJ. Visual estimation versus quantitative assessment of left ventricular ejection fraction: a comparison by cardiovascular magnetic resonance imaging. Am Heart J 2005; 150: 737-42. [Crossref]

12. Holloway CJ, Edwards LM, Rider OJ, Fast A, Clarke K, Francis JM, et al. A comparison of visual and quantitative assessment of left ventricular ejection fraction by cardiac magnetic resonance. Int J Cardiovasc Imaging 2011; 27: 563-9. [Crossref]

13. McGowan JH, Cleland JG. Reliability of reporting left ventricular systolic function by echocardiography: a systematic review of 3 methods. Am Heart J 2003; 146: 388-97. [Crossref]

14. Blondheim DS, Beeri R, Feinberg MS, Vaturi M, Shimi S, Fehske W, et al. Reliability of visual assessment of global and segmental left ventricular function: a multicenter study by the Israeli Echocardiography Research Group. J Am Soc Echocardiogr 2010; 23: 258-64. [Crossref]

15. Kusunose K, Shibayama K, Iwano H, Izumo M, Kagiyama N, Kurosawa K, et al.; JAYEF Investigators. Reduced variability of visual left ventricular ejection fraction assessment with reference images: The Japanese Association of Young Echocardiography Fellows multicenter study. J Cardiol 2018; 72: 74-80. [Crossref]