Original article

Test–re-test reproducibility of Doppler echocardiography for assessment of electromechanical dyssynchrony: Implications for heart failure clinic

Vittorio Palmieri (MD, PhD)a,*, Cesare Russo (MD)b, Antonietta Buonomo (MD)a, Giovanvirgilio Cimmino (MD)a, Donato Tartaglione (MD)a, Salvatore Pezzullo (MD)c, Aldo Celentano (MD)a

a Heart Failure Unit, "Unità complessa di Cardiologia, Ospedale dei Pellegrini – ASL Napoli 1", Naples, Via Napoli 816, San Felice a Cancello (CE) 81027, Italy
b Department of Medicine, Columbia University, New York, NY, USA
c Department of Clinical and Experimental Medicine, "Federico II" University Hospital, Naples, Italy

Received 20 February 2010; received in revised form 3 May 2010; accepted 10 June 2010
Available online 16 July 2010

KEYWORDS
Heart failure; Echocardiography; Transthoracic; Dyssynchrony; Methodology

Summary
Background: Reproducibility of Doppler echocardiography for assessment of inter-ventricular and intra-left ventricular (LV) dyssynchrony, and its clinical implications, have not been established.
Methods: Twenty-eight subjects (heart failure stages A–C, 61% with QRS ≥ 120 ms, ejection fraction (EF) ≤35%) underwent two consecutive echo-studies within 24 h to evaluate test–re-test reproducibility of inter-ventricular electromechanical delay (VV delay, by traditional pulsed-Doppler), and intra-LV electromechanical delay between opposite LV walls by color-coded Doppler tissue-velocity (COLOR-DTI), and by pulsed-Doppler tissue spectrum (PW-DTI). Reproducibility of LV internal diastolic diameter (LVIDD) and of EF (by Simpson’s method) assessments was evaluated contextually for reference.
Results: Intra-study and inter-study reproducibility of inter-ventricular and intra-LV electromechanical dyssynchrony was in general good, and comparable to the reproducibility of LVIDD and EF assessments. Between-study reproducibility of PW-TDI method was fair, but showed poor agreement with COLOR-TDI method. In repeated studies, agreement of significant electromechanical delay by COLOR-TDI was comparable...
to the agreement of EF ≤ 35%. In the 5 patients who had simultaneously large QRS, EF ≤ 35%, and significant inter- and intra-ventricular dyssynchrony at study #1, 3 had EF 36–40% and 1 showed no significant dyssynchrony by study #2.

**Conclusion:** In serial echocardiographic studies, Doppler echocardiography showed a good test–re-test reproducibility for the identification of significant electromechanical delay. Planimetry for EF assessment was a source of variability as relevant as Doppler echocardiography, but COLOR-DTI may add meaningful and reproducible information to QRS duration for cardiac-resynchronization therapy.

© 2010 Japanese College of Cardiology. Published by Elsevier Ireland Ltd. All rights reserved.

---

**Introduction**

Echocardiography allows evaluations of inter-ventricular and intra-left ventricular (LV) electromechanical dyssynchrony [1–3], therapeutical targets of cardiac-resynchronization therapy in congestive heart failure [4–16]. However, echocardiographic assessment of electromechanical dyssynchrony is not recommended for screening congestive heart failure patients who are candidates for cardiac-resynchronization therapy [17], at least in part because echocardiography is not considered sufficiently reproducible [18].

However, acute reduction of intra-LV electromechanical delay by biventricular pacing [15], and pacing at LV site of maximum electromechanical delay, is both predictors of response to cardiac-resynchronization therapy [19]. Moreover, echocardiography contributes to the evaluation of nonresponders to cardiac-resynchronization therapy [20]. Thus, standard clinical practice relies significantly on echocardiographic assessment of electromechanical dyssynchrony in congestive heart failure patients.

In the Predictors of Response to Cardio-resynchronization Therapy (PROSPECT) multicenter non-randomized trial study, variability of echocardiographic parameters of electromechanical dyssynchrony between-readings was elevated, and considered one of the main reasons for the lack of capability of echocardiography to identify responders to cardiac-resynchronization therapy [18]. However, test–re-test reproducibility of electromechanical dyssynchrony was not evaluated in the PROSPECT. The latter is an important issue because in common clinical practice, patients with congestive heart failure are evaluated in single centers for serial assessments of a number of parameters for selection of candidates for cardiac-resynchronization therapy, such as LV ejection fraction (EF) and LV internal diameter, and other parameters useful to explore residual inter- or intra-LV dyssynchrony in nonresponders [20]. To this extent, knowledge of the confidence intervals of echocardiographic parameters of electromechanical dyssynchrony assessed in repeated studies (i.e. test–re-test variability) is an important clinical tool to interpret differences between consecutive evaluations. In the present study, we tested the hypothesis that the reproducibility of echocardiography for the assessment of traditional Doppler and tissue Doppler-based parameters of cardiac dyssynchrony in serial echocardiographic evaluations (i.e. inter-study, intra-subjects, intra-reader, and between-reader test–re-test reproducibility) may be close to the variability observed in consecutive readings of single studies (i.e. intra-study, intra-subjects test–re-test reproducibility); furthermore, we also tested the hypothesis that reproducibility of echocardiography for the assessment of dyssynchrony may be as good as reproducibility of echocardiographic assessment LV chamber size and EF, mandatory for screening and follow-up studies in patients with congestive heart failure.

**Methods**

**Design and patients**

According to previous studies [21–23], we designed a protocol to analyze biological and technical sources of variability in patients (in-hospital or ambulatory, at the Cardiology Unit, “Ospedale dei Pellegrini, Azienda Sanitaria Locale-Napoli 1”, Naples, Italy), willing to undergo two echocardiographic studies 24 h apart. All subjects had indications for echocardiographic study because of stage A or C congestive heart failure, and gave informed consent to enter the study. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki as reflected in *a priori* approval by the local institution’s human research committee. Because truncated distributions of measurements may influence variability for repeated measurements, the study sample included clinically healthy subjects as well as patients with chronic congestive heart failure in order to obtain the largest distribution of possible measurements for parameters of LV size, systolic function, and dyssynchrony, also considering that patients with heart failure may have normal LVEF, and that LV dyssynchrony may be present in the absence of large QRS as well as in those without severely depressed LVEF [24]. According to a pre-defined protocol, a single operator performed two studies per subject within 24 h, using the same echocardiographic machine and a standardized echocardiographic protocol. A pre-specified and standardized reading protocol was also applied. The first reader evaluated the first study immediately, the second study in a random sequence 1–3 months later, and the first study was re-evaluated in random sequence after an additional 1–3 months. Therefore, by study protocol, 3 readings were available for each parameter: readings of the study #1 (R1), re-readings of the study #1 (re-R1) and readings of the study #2 (R2). Thus, the dataset comprised 84 readings (R1, re-R1, and R2) obtained from 56 echocardiographic studies in 28 participants. In addition, parts of R1 data were re-evaluated by a second reader to test between-reader intra-study variability. Treatment regimen remained stable between the two echocardiographic studies; however, changes in clinical conditions requiring treatment regimen modification represented an exclusion criterion. In fact,
heart rate (68 ± 5 bpm vs. 68 ± 6 bpm) and blood pressure did not differ between R1 and R2 (121/74 ± 21/11 mm Hg vs. 120/78 ± 21/9 mm Hg) (both p > 0.5). In addition, we excluded a priori subjects with atrial fibrillation, and those with frequent ventricular or supraventricular ectopic beats.

Echocardiographic procedures

Echocardiographic studies were performed by a commercially available echocardiographic machine (GE VIVID 7, GE, Horten, Norway, BT 6.0.1), using a broadband transducer (MS4), and second harmonic modality 1.7–3.4 MHz or 1.9–3.8 MHz when needed applying current standards [1—3,21–23,25,26]. Digital loops of 4 cardiac cycles were stored while patients held their breath.

Traditional echocardiographic methodology

LV internal diastolic diameter (LVIDD), LVEF by Simpson’s method, transaortic, and trans-pulmonary flows by traditional pulsed-Doppler was evaluated by standard methods [22,23,25,27]. LV and right ventricular (RV) pre-ejection times were measured as the time between QRS-onset and trans-valvular flow onset; low-velocity reject filter was appropriately reduced for optimal visualization of Doppler spectrum.

Pulsed-Doppler issue imaging

Pulsed-Doppler tissue interrogation (PW-DTI, pulse repetition frequency ≥250 Hz) was used to evaluate myocardial velocities by placing a 6-mm wide sample volume on the mitral anulus at the 4 LV corners (inferior septum, anterolateral in apical 4 chamber view, anterior and inferior in apical 2 chamber view), minimizing the angle between the Doppler cursor beam and the longitudinal direction of the mitral annular motion.

Color-coded tissue-velocity imaging

Two-dimensional echocardiography with over-imposed color-coded myocardial velocity map (COLOR-DTI) was performed; sector size, depth, and two-dimensional gain were adjusted to obtain highest possible pulse repetition frequency and frame rate, and best color saturation avoiding tissue-velocity aliasing (frame rate between 90 and >200 frames/s; pulse repetition frequencies between 500 Hz and 1 kHz, aliasing velocities between 12 and 32 cm/s). COLOR-DTI aliasing velocity was set as ±20 cm/s for patients with normal LVEF, and reduced to ±16 or ±12 cm/s for those with moderately or severely reduced EF.

Inter-ventricular dyssynchrony

Inter-ventricular electromechanical delay (VV delay) was calculated as the difference between LV and RV pre-ejection times.

Intra-LV dyssynchrony

M-mode based methodology to assess intra-LV dyssynchrony was not evaluated because it was previously demonstrated that such a method is feasible in approximately 60% of the candidates to cardio-resynchronization therapy [17,28]. By PW-TDI, myocardial-systolic-wave (S) was identified with reference to the QRS as the first positive wave peaking after aortic valve opening using LV pre-ejection time for reference. Electromechanical activation of each segment was defined as the time between QRS-onset and the S onset (Ts-onset). Acquisition of time-to-peak S by PW-DTI was initially planned and subsequently aborted after an interim analysis on the first 5 cases revealed excessive intra- and inter-study variability, consistent with other reports [2,29]. Thus, by PW-DTI, Ts-onset (septal-anterolateral) measured electromechanical delay between anterolateral wall and inferior inter-ventricular septum, and Ts-onset (inferior—anterior) measured the electromechanical delay between the anterior and the inferior walls, with greater values indicating greater electromechanical delay between opposite walls [6]. Maximum electromechanical delay at basal LV segments (longest—shortest Ts-onset [5]) was assessed only by PW-DTI because it was demonstrated previously that COLOR-DTI method is not reproducible for the assessment of such a parameter [30]. By COLOR-DTI, a 4-basal segments approach required offline reconstruction of the myocardial velocity curves. Preliminarily, transaortic flow was used to mark LV ejection phase. A 4-mm wide and 8-mm long sample volume was placed at the basal portion of the inferior septum and of the anterolateral wall in end-systole (marker of aortic valve closure), using manual tracking function for back-repositioning the region of interest on the basal segment at end-diastole [31]. Electromechanical delay between opposite walls was measured as the time elapsing between the peak systolic velocities, i.e. Ts-peak (septa-antrolateral) and Ts-peak (inferior— anterior), with greater values indicating greater electromechanical delay between opposite walls. Identification of the reference regional peak systolic velocity was based on the highest velocity obtained during LV ejection phase; eventually, controversies were resolved by converting TDI curves in regional longitudinal strain to exclude systolic peak due to tracking rather than to longitudinal contractility.

Secondly, the standard deviation of the time-to-peak myocardial systolic velocity from 12 segments explored (Ts-SD), 6 basal and 6 at mid-cavity segments in apical 4, 2
Statistical analysis

Analysis of variance for repeated measurements was used to evaluate between-study differences. Intraclass correlation coefficients of reliability and their 95% confidence interval [32] were estimated using SPSS for Windows (release 9.0.1, February 1999, Copyright SPSS Inc., Chicago, IL, USA), according to a two-way random model for absolute agreement, as in previous similar studies [21,22]. Higher intraclass correlation coefficient values indicate higher reproducibility of measurements. Systematic errors across the range of measurements were explored according to procedures described by Bland—Altman [33] reporting confidence intervals of between-reading differences estimated at 90% and 80%. In Table 3, the proportion of subjects classified as having inter- and intra-ventricular dyssynchrony using data from the three sets of readings are reported with reference to R1. The Cohen’s kappa statistic was performed as measure of between-study agreement. Higher kappa indicated higher between-study agreement. For all tests, null-hypothesis was rejected for two-tailed \( p < 0.05 \).

Results

Study sample characteristics

Main characteristics of the study sample are reported in Table 1. Among the 56 studies obtained (2 per subject), technical quality was good in 43%, sufficient or just sufficient in 57%; the rate of exam rejection (due to poor quality) was approximately 8%.

In preliminary analyses, QRS duration correlated with VV delay \((r = 0.69, p < 0.01)\), with EF \((r = -0.66, p < 0.01)\), and with Ts-SD \((r = 0.50, p < 0.05)\), but not with Ts-peak and Ts-onset (all \(-0.3 < r < 0.3, \) all \( p > 0.2 \)). EF correlated inversely with VV delay \((r = -0.40, p < 0.01)\) and with Ts-SD \((r = 0.39, p < 0.05)\), but not with Ts-peak and Ts-onset (all \(-0.3 < r < 0.3, \) all \( p > 0.05 \)). VV delay correlated moderately with Ts-SD \((r = 0.51, p < 0.01)\), but not with Ts-peak and Ts-onset (all \(-0.3 < r < 0.3, \) all \( p > 0.1 \)).

1.1. Test–re-test reproducibility, intra-reader, intra- and between-study, according to intraclass correlation coefficients

Intra-study (between-readings) and between-study test–re-test reproducibility of parameters of electromechanical dyssynchrony was between fair and good (Table 2), and

<table>
<thead>
<tr>
<th>Parameter</th>
<th>R1</th>
<th>re-R1</th>
<th>R2</th>
<th>Intraclass correlation coefficients and their 95% confidence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Ts-SD, ms</td>
<td>45 ± 22</td>
<td>42 ± 19</td>
<td>44 ± 23</td>
<td>0.90 (0.80–0.94) 0.88 (0.75–0.95)</td>
</tr>
<tr>
<td>Ts-peak (anterolateral–inferior septal), ms</td>
<td>54 ± 81</td>
<td>60 ± 91</td>
<td>53 ± 90</td>
<td>0.94 (0.89–0.99) 0.97 (0.93–0.98)</td>
</tr>
<tr>
<td>Ts-onset (anterolateral–inferior septal), ms</td>
<td>-6 ± 42</td>
<td>0 ± 41</td>
<td>-22 ± 38</td>
<td>0.91 (0.81–0.98) 0.63 (0.31–0.82)</td>
</tr>
<tr>
<td>Ts-peak (anterior–inferior), ms</td>
<td>-17 ± 105</td>
<td>-14 ± 120</td>
<td>-19 ± 111</td>
<td>0.98 (0.97–0.99) 0.97 (0.93–0.98)</td>
</tr>
<tr>
<td>Ts-onset (anterior–inferior), ms</td>
<td>24 ± 27</td>
<td>19 ± 36</td>
<td>25 ± 29</td>
<td>0.73 (0.49–0.87) 0.70 (0.44–0.86)</td>
</tr>
<tr>
<td>LV pre-ejection time, ms</td>
<td>102 ± 37</td>
<td>101 ± 36</td>
<td>102 ± 38</td>
<td>0.98 (0.95–0.99) 0.97 (0.94–0.99)</td>
</tr>
<tr>
<td>RV pre-ejection time, ms</td>
<td>73 ± 22</td>
<td>74 ± 20</td>
<td>72 ± 20</td>
<td>0.93 (0.85–0.97) 0.93 (0.86–0.97)</td>
</tr>
<tr>
<td>VV delay, ms</td>
<td>31 ± 26</td>
<td>28 ± 27</td>
<td>30 ± 27</td>
<td>0.91 (0.83–0.96) 0.90 (0.79–0.95)</td>
</tr>
<tr>
<td>LVIDD, cm</td>
<td>6.3 ± 1.0</td>
<td>6.2 ± 1.0</td>
<td>6.1 ± 1.0</td>
<td>0.97 (0.94–0.98) 0.96 (0.94–0.99)</td>
</tr>
<tr>
<td>EF, %</td>
<td>40 ± 18</td>
<td>42 ± 18</td>
<td>42 ± 19</td>
<td>0.94 (0.88–0.97) 0.95 (0.89–0.98)</td>
</tr>
</tbody>
</table>
was close to that observed for repeated assessments of LVDD and EF. However, on average, reproducibility was superior for COLOR-DTI and PW-DTI parameters. In particular, reproducibility of TS-onset (septal-anterolateral) and TS-onset (inferior—anterolateral) studies was poorer compared to between-readings intra-study (all \( p < 0.05 \)). Assessment of LV pre-ejection time was more reproducible than that of RV pre-ejection time. Reproducibility of assessment of maximum LV basal electromechanical delay by PW-DTI was relatively good in intra-study (4 ± 38 ms in R1 vs. 6 ± 41 ms in re-R1, intraclass correlation coefficient and its 95% confidence interval [0.83, 0.66–0.94], but was suboptimal in between-study (−12 ± 37 ms in R2, intraclass correlation coefficient and its 95% confidence interval 0.60, 0.22–0.75).

### Confidence limits of intra-reader, between-study and intra-study variability

Bland–Altman method excluded systematic errors in reading procedures (all Pearson correlation coefficients comprised between −0.2 and 0.2, all \( p > 0.5 \), data not shown). In Table 3, it may be appreciated that 80% and 90% confidence intervals of the differences reading-by-reading were slightly wider in between-study than in intra-study.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Re-R1—R1 mean difference (5th, 10th, 90th, 95th percentile of the difference between reading intra-studies)</th>
<th>R2—R1 mean difference (5th, 10th, 90th, 95th percentile of the difference between studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS-SD, ms</td>
<td>−3 ± 10 (−24; −16; +9; +13)</td>
<td>−2 ± 10 (−18; −14; +13; +17)</td>
</tr>
<tr>
<td>TS-peak (anterolateral–inferior septal), ms</td>
<td>7 ± 30 (−62; −31; +40; +60)</td>
<td>−1 ± 23 (−45; −31; +32; +50)</td>
</tr>
<tr>
<td>TS-onset (anterolateral–inferior septal), ms</td>
<td>6 ± 17 (−25; −12; +32; +36)</td>
<td>−16 ± 32 (−98; −71; +71; +38)</td>
</tr>
<tr>
<td>TS-peak (anterior–inferior), ms</td>
<td>3 ± 18 (−26; −20; +21; +36)</td>
<td>−2 ± 29 (−40; −31; +21; +74)</td>
</tr>
<tr>
<td>TS-onset (anterior–inferior), ms</td>
<td>−5 ± 18 (−76; −33; +16; +21)</td>
<td>1 ± 22 (−58; −24; +25; +29)</td>
</tr>
<tr>
<td>LV pre-ejection time, ms</td>
<td>−1 ± 8 (−11; −8; +9; +18)</td>
<td>−1 ± 9 (−17; −10; +10; +16)</td>
</tr>
<tr>
<td>RV pre-ejection time, ms</td>
<td>2 ± 7 (−10; −8; +11; +19)</td>
<td>0 ± 8 (−11; −8; +10; +19)</td>
</tr>
<tr>
<td>VV delay, ms</td>
<td>−3 ± 10 (−26; −14; +14; +17)</td>
<td>−1 ± 12 (−30; −15; +15; +17)</td>
</tr>
<tr>
<td>LVIDD, cm</td>
<td>0.0 ± 0.2 (−0.4; −0.3; +0.3; +0.4)</td>
<td>0.0 ± 0.2 (−0.5; −0.4; +0.2; +0.3)</td>
</tr>
<tr>
<td>EF, %</td>
<td>2 ± 6 (−7; −6; +7; +13)</td>
<td>2 ± 6 (−8; −7; +11; +12)</td>
</tr>
</tbody>
</table>

Data in table are mean ± standard deviation of between-readings differences, and limits of 90% and 80% confidence intervals of between-readings difference. Abbreviations as in Table 2.

### Reclassification rates of significant electromechanical delay due to observed variability

As reported in Table 4, for each parameter, agreement between R1 data re-R1 and R2 data was relatively good. Those data were consistent among patients with QRS > 120 ms and EF ≤ 35% (data not shown). However, agreement between TS-onset and TS-peak (septal-anterolateral) was suboptimal (Fig. 1, intraclass correlation coefficient = 0.04, 95% confidence interval −0.23, +0.36; in fact, of the patients with TS-peak (septal-anterolateral) >60 ms, only one had TS-onset (septal-anterolateral) >60 ms, whereas of the patients with TS-peak (septal-anterolateral) ≤60 ms, 100% also had TS-onset (septal-anterolateral) ≤60 ms (Kappa = 0.1, \( p = N.S. \)). Agreement between TS-SD > 33 ms and TS-peak (septal-anterolateral) >60 ms was also poor (Kappa = 0.19, \( p = N.S. \)).

### Simulation of the impact of reproducibility on screening for cardiac-resynchronization therapy

Seventeen subjects showed large QRS and EF ≤35% by R1 (Fig. 2A). Of those, EF by R2 yielded a reclassified rate of 29% (in all cases comprised between 36% and 44% by R2); a sig-

---

Table 3: Between-reading differences of echocardiographic parameters and their limits at 80% and 90% confidence intervals.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Re-R1—R1 mean difference (5th, 10th, 90th, 95th percentile of the difference between reading intra-studies)</th>
<th>R2—R1 mean difference (5th, 10th, 90th, 95th percentile of the difference between studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS-SD, ms</td>
<td>−3 ± 10 (−24; −16; +9; +13)</td>
<td>−2 ± 10 (−18; −14; +13; +17)</td>
</tr>
<tr>
<td>TS-peak (anterolateral–inferior septal), ms</td>
<td>7 ± 30 (−62; −31; +40; +60)</td>
<td>−1 ± 23 (−45; −31; +32; +50)</td>
</tr>
<tr>
<td>TS-onset (anterolateral–inferior septal), ms</td>
<td>6 ± 17 (−25; −12; +32; +36)</td>
<td>−16 ± 32 (−98; −71; +71; +38)</td>
</tr>
<tr>
<td>TS-peak (anterior–inferior), ms</td>
<td>3 ± 18 (−26; −20; +21; +36)</td>
<td>−2 ± 29 (−40; −31; +21; +74)</td>
</tr>
<tr>
<td>TS-onset (anterior–inferior), ms</td>
<td>−5 ± 18 (−76; −33; +16; +21)</td>
<td>1 ± 22 (−58; −24; +25; +29)</td>
</tr>
<tr>
<td>LV pre-ejection time, ms</td>
<td>−1 ± 8 (−11; −8; +9; +18)</td>
<td>−1 ± 9 (−17; −10; +10; +16)</td>
</tr>
<tr>
<td>RV pre-ejection time, ms</td>
<td>2 ± 7 (−10; −8; +11; +19)</td>
<td>0 ± 8 (−11; −8; +10; +19)</td>
</tr>
<tr>
<td>VV delay, ms</td>
<td>−3 ± 10 (−26; −14; +14; +17)</td>
<td>−1 ± 12 (−30; −15; +15; +17)</td>
</tr>
<tr>
<td>LVIDD, cm</td>
<td>0.0 ± 0.2 (−0.4; −0.3; +0.3; +0.4)</td>
<td>0.0 ± 0.2 (−0.5; −0.4; +0.2; +0.3)</td>
</tr>
<tr>
<td>EF, %</td>
<td>2 ± 6 (−7; −6; +7; +13)</td>
<td>2 ± 6 (−8; −7; +11; +12)</td>
</tr>
</tbody>
</table>

Data in table are mean ± standard deviation of between-readings differences, and limits of 90% and 80% confidence intervals of between-readings difference. Abbreviations as in Table 2.

Table 4: Between-readings agreement of parameters of inter- and intra-ventricular dyssynchrony.

<table>
<thead>
<tr>
<th>Given data by R1</th>
<th>VV delay</th>
<th>TS-SD</th>
<th>TS-peak (anterolateral–inferior septal)</th>
<th>TS-onset (anterolateral–inferior septal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cut-points ms</td>
<td>&gt;40 ≤40</td>
<td>&gt;33 ≤33</td>
<td>&gt;60 ≤60</td>
<td>&gt;60 ≤60</td>
</tr>
<tr>
<td>Confirmed by re-R1</td>
<td>78% 94%</td>
<td>85% 88%</td>
<td>93% 92%</td>
<td>100% 100%</td>
</tr>
<tr>
<td>Kappa (p)</td>
<td>0.74 (&lt;0.001)</td>
<td>0.67 (&lt;0.005)</td>
<td>0.86 (&lt;0.001)</td>
<td>1 (0.000)</td>
</tr>
<tr>
<td>Confirmed by R2</td>
<td>78% 94%</td>
<td>95% 88%</td>
<td>80% 100%</td>
<td>100% 100%</td>
</tr>
<tr>
<td>Kappa (p)</td>
<td>0.74 (&lt;0.001)</td>
<td>0.83 (&lt;0.001)</td>
<td>0.79 (0.002)</td>
<td>1 (0.000)</td>
</tr>
</tbody>
</table>

Abbreviations as in Table 2.
significant septal-anterolateral electromechanical delay was found in approximately 71% by R1, confirmed in 77% by re-R1, and in 59% by R2; similarly, a significant inter-ventricular dyssynchrony was found in 47% by R1, re-R1, and R2 data.

By a different approach, 12 subjects were selected by large QRS and T<sub>s-peak</sub> (septal-anterolateral) >60 ms (Fig. 2B); in those, T<sub>s-peak</sub> (septal-anterolateral) >60 ms was not confirmed in 17% by R2. EF, which was ≤35% by R1 in 100% of the selected 12 subjects, and confirmed in 92% by re-R1, was between 36% and 44% by R2 in approximately 5/12. Interestingly, significant VV delay was found in approximately 4/12 of the subjects pre-selected by large QRS and T<sub>s-peak</sub> (septal-anterolateral) >60 ms.

With more stringent selection criteria applied, i.e. QRS ≥120 ms, EF ≤35%, VV delay >40 ms, and T<sub>s-peak</sub> (anterolateral—inferior septal) >60 ms at R1, 5 subjects were selected; in those, EF was between 36% and 45% in 3/5 by R2, whereas the rate of reclassification was 0% for QRS ≥120 ms, 1/5 for VV delay >40 ms, and 1/5 for T<sub>s-peak</sub> (anterolateral—inferior septal) >60 ms (Fig. 3).

**Between-reader variability in assessment of inter-and intra-ventricular dyssynchrony**

Giving results of intra-reader test—re-test reproducibility, the focus of set of re-readings by a second reader was limited to VV delay, T<sub>s-peak</sub> (anterolateral-inferior septal), LVIDD, and EF. Intraclass correlation coefficients were: 0.86 (0.72–0.93) for VV delay; 0.92 (0.84–0.96) for T<sub>s-peak</sub> (anterolateral-inferior septal); 0.82 (0.64–0.91) for LVIDD; 0.89 (0.78–0.95) for EF. Mean, standard deviations and 80% confidence intervals of between-reader difference were: for VV delay: −1 ± 13 ms, −16/+18 ms; for T<sub>s-peak</sub> (anterolateral-inferior septal): 8 ± 30 ms, −31/+51 ms; for LVIDD: −0.1 ± 0.6 cm, −0.3/+0.4 cm; for EF: 3 ± 8%, −8/+11%.
Test—re-test reproducibility of Doppler echocardiography for assessment of electromechanical dyssynchrony: 277

approach may actually reduce drastically the number of QRS duration (Fig. 3), suggesting that a multi-parametric when combined with Doppler echocardiography and large relevant source of variability for patient classification even more relevant than Ts-peak (septal-anterolateral) has the potential to add useful information for patient selection for cardiac-resynchronization therapy. Response to cardiac-resynchronization therapy among patients with QRS between 120 and 149 ms and EF ≤ 35% is less convincing than in those with QRS ≥ 150 ms [8], and cardiac-resynchronization therapy appears to be effective for the prevention of symptomatic congestive heart failure [20]. In our study, large QRS duration essentially predicted depressed EF and inter-ventricular dyssynchrony; in addition, longer VV delay tended to predict global intra-LV dyssynchrony (i.e. wider Ts-SD) but not specific intra-LV targets for resynchronization therapy, as expected [36—38]. In clinical practice, measures of intra-LV dyssynchrony are useful for identification of crucial targets for cardiac-resynchronization therapy [19], measuring the efficacy of resynchronization [15], and re-assessment of nonresponders [20]. To such an extent, therefore, test—re-test reproducibility of echocardiography is a key factor. Fig. 2A and B suggests that in terms of reliable patient classification, combination of large QRS with depressed EF appears to be potentially less effective than combination of large QRS with significant intra-LV dyssynchrony. In fact, in the case of combination of large QRS and depressed EF, of the 17 subjects identified, 29% had EF >35% by re-evaluation, but only approximately 2/3 showed significant intra-LV target for resynchronization therapy. In a different scenario, combination of large QRS with Ts-peak (septal-anterolateral) >60 ms identified 12 subjects (5 less than in the previous scenario); in those, significant intra-LV target for resynchronization therapy was confirmed in 83% by R2. Interestingly, of those 12 subjects, more than 90% also had EF ≤ 35% by R1 and re-R1 data, whereas when EF was re-evaluated, it was confirmed to be ≤35% in approximately 60%, and EF was between 36% and 44% in the remaining 40%. Considering that random fluctuation of test—re-test estimation of EF by R2 ranged between −7% and +11% from any given value obtained in R1 (Table 3), findings shown in Tables 2 and 3 are not surprising, and suggest that in “real world” subjects with EF = 30% by R1 may show an EF value ranging between 23% and 41% by a re-evaluation within 24 h due to variability of the method. Our data are additive to those reported in the PROSPECT, because our data are referred to test—re-test evaluation and not only to re-readings of a single study. A future challenge would be to evaluate which strategy may be more cost-effective, the one that might rely on the use of more stringent cut-points (EF ≤ 30% and QRS > 130 ms, for instance) to potentially reduce the impact of the variability of the method for patient selection and risk not to deliver an appropriate therapy to a significant proportion of patients, or in alternative a second one considering combination of large QRS with specific identification of targets for cardiac-resynchronization therapy (i.e. measures of significant electromechanical delay) to propose cardiac-resynchronization therapy. Another practical aspect touched on by our study is the one related to the use of echocardiography for serial assessment of patients and/or for device optimization. When groups of patients are compared, even relatively small between-group mean differences may be statistically significant with large sample size and small standard deviations. However, in standard clinical practice, data shown in Table 3 are a useful guide because they suggest that VV delay assessed in a single subject in two studies may range between −15 and +15 ms at 80% confidence interval independent to operators. The finding implies that a difference in VV delay between two consecutive studies

Figure 3 Rate of reclassification by R2 of subjects who by R1 had QRS duration ≥120 ms, EF ≤ 35%, VV delay >40 ms, and Ts-peak (anterolateral-inferior septal) >60 ms. Gray bars represent % of reclassification; the absence of the bar associated with QRS <120 ms is due to no reclassification of large QRS between R1 and R2 in the subgroup. R1, data from first reading of the first exam; R2, data from readings of the second exam; Ts-peak, time between the peak of the myocardial systolic wave by color-coded tissue Doppler imaging at the basal myocardial segments of opposite walls; EF, left ventricular ejection fraction by Simpson’s method.

Discussion

Our study showed that the reproducibility of Doppler echocardiography for the assessment of electromechanical dyssynchrony can be fairly good for parameters such as VV delay, Ts-SD, and Ts-peak (septal-anterolateral), and substantially comparable to that observed for serial assessments of LVIDD and of EF. This implies that standardization of echocardiographic procedures may reduce variability between readers, as shown previously for the assessment of LV mass and for Doppler parameters of LV diastolic function [21,34], which may have implications for heart failure clinics.

In fact, findings summarized in Fig. 2A and B showed that in single patients with large QRS, assessment of EF for patients’ classification may be a source of variability more relevant than Ts-peak (septal-anterolateral), which identifies possible targets for resynchronization therapy. Moreover, planimetry for EF assessment remained a relevant source of variability for patient classification even when combined with Doppler echocardiography and large QRS duration (Fig. 3), suggesting that a multi-parametric approach may actually reduce drastically the number of candidates to cardiac-resynchronization therapy [35], without improving the reliability of the method for patients’ classification.

Our results also reinforce the notion that Ts-peak (septal-anterolateral) has the potential to add useful information for patient selection for cardiac-resynchronization therapy. Response to cardiac-resynchronization therapy among patients with QRS between 120 and 149 ms and EF ≤ 35% is less convincing than in those with QRS ≥ 150 ms [8], and cardiac-resynchronization therapy appears to be effective for the prevention of symptomatic congestive heart failure [20]. In our study, large QRS duration essentially predicted depressed EF and inter-ventricular dyssynchrony; in addition, longer VV delay tended to predict global intra-LV dyssynchrony (i.e. wider Ts-SD) but not specific intra-LV targets for resynchronization therapy, as expected [36—38]. In clinical practice, measures of intra-LV dyssynchrony are useful for identification of crucial targets for cardiac-resynchronization therapy [19], measuring the efficacy of resynchronization [15], and re-assessment of nonresponders [20]. To such an extent, therefore, test—re-test reproducibility of echocardiography is a key factor. Fig. 2A and B suggests that in terms of reliable patient classification, combination of large QRS with depressed EF appears to be potentially less effective than combination of large QRS with significant intra-LV dyssynchrony. In fact, in the case of combination of large QRS and depressed EF, of the 17 subjects identified, 29% had EF >35% by re-evaluation, but only approximately 2/3 showed significant intra-LV target for resynchronization therapy. In a different scenario, combination of large QRS with Ts-peak (septal-anterolateral) >60 ms identified 12 subjects (5 less than in the previous scenario); in those, significant intra-LV target for resynchronization therapy was confirmed in 83% by R2. Interestingly, of those 12 subjects, more than 90% also had EF ≤ 35% by R1 and re-R1 data, whereas when EF was re-evaluated, it was confirmed to be ≤35% in approximately 60%, and EF was between 36% and 44% in the remaining 40%. Considering that random fluctuation of test—re-test estimation of EF by R2 ranged between −7% and +11% from any given value obtained in R1 (Table 3), findings shown in Tables 2 and 3 are not surprising, and suggest that in “real world” subjects with EF = 30% by R1 may show an EF value ranging between 23% and 41% by a re-evaluation within 24 h due to variability of the method. Our data are additive to those reported in the PROSPECT, because our data are referred to test—re-test evaluation and not only to re-readings of a single study. A future challenge would be to evaluate which strategy may be more cost-effective, the one that might rely on the use of more stringent cut-points (EF ≤ 30% and QRS > 130 ms, for instance) to potentially reduce the impact of the variability of the method for patient selection and risk not to deliver an appropriate therapy to a significant proportion of patients, or in alternative a second one considering combination of large QRS with specific identification of targets for cardiac-resynchronization therapy (i.e. measures of significant electromechanical delay) to propose cardiac-resynchronization therapy. Another practical aspect touched on by our study is the one related to the use of echocardiography for serial assessment of patients and/or for device optimization. When groups of patients are compared, even relatively small between-group mean differences may be statistically significant with large sample size and small standard deviations. However, in standard clinical practice, data shown in Table 3 are a useful guide because they suggest that VV delay assessed in a single subject in two studies may range between −15 and +15 ms at 80% confidence interval independent to operators. The finding implies that a difference in VV delay between two consecutive studies...
should exceed 15 ms (in either direction) in order to have less than 10% likelihood to be a random effect as opposed to a therapeutic effect. Similarly, the range of variability between two consecutive assessments of EF in a single subject may range between −7% and +11% at 80% confidence interval intra-reader. Thus, a patient with EF = 35% and a VV delay of 40 ms may show an EF between 28% and 46% and a VV delay between 25 and 60 ms in a second evaluation within 24 h just by chance with a confidence interval of 80%. This may be important when patients are re-evaluated for lack of response to cardiac-resynchronization therapy, or for device optimization.

**Study limitations**

The present study is based on data from a single center, and single technology, which may be considered “‘best scenario”. However, our clinical scenario reflects classic clinical practice, and may be used for reference. Our study is based on 28 subjects selected not consecutively but randomly, which may be considered a small population sample; nevertheless, we actually obtained 56 studies, and 3 sets of readings (i.e. 84 readings) from a first reader, and additional 28 readings from a second reader; thus, to the best of our knowledge, this is the largest dataset in which test−re-test reproducibility of Doppler echocardiography for LV dyssynchrony has been evaluated. In our study, reading procedures were standardized and pre-specified as a fundamental part of the echocardiographic protocol, as in previous studies [21–23,34]; it is likely that such an approach contributed to reducing intra-reader between-reading and between-reader variability, such as in the case of multiple peak systolic velocity in myocardial regions detected by COLOR-DTI modality; however, whether our method yielded good test−re-test reproducibility may also result in greater reliability of COLOR-TDI in identifying responders to cardiac-resynchronization therapy needs to be tested. Nevertheless, it is worth noting that when standardized methodology is used to identify significant intra-ventricular dyssynchrony and acute response to resynchronization therapy, COLOR-TDI was found to be highly reliable in predicting responders to cardiac-resynchronization therapy [15]. Our study included patients with normal EF and short QRS as well as those with depressed EF and large QRS duration to avoid bias introduced by truncated statistical distributions. The present study did not examine speckle-tracking data for assessment of novel parameters of electromechanical dyssynchrony [39], which are much less available in common echo-laboratories compared to traditional and COLOR-DTI.

**Conclusions**

Reproducibility of echocardiography for serial assessments of inter-ventricular and intra-LV dyssynchrony may be good and comparable to that for serial assessments of LV size and EF by planimetry. Combination of prolonged QRS duration with depressed EF and significant cardiac electromechanical dyssynchrony did not limit the impact of variability of EF on patients’ classification in serial studies, but reduced significantly the number of possible candidates to resynchronization therapy. In contrast, combination of QRS duration with significant intra-LV dyssynchrony showed good test−re-test reproducibility, identified a large part of subjects with significantly reduced EF, and did not reduce excessively the number of patients who were candidates for cardiac-resynchronization therapy. Future studies are needed to evaluate which strategy for patient selection for cardiac-resynchronization therapy may be more cost-effective.

**References**

imaging accurately measures left ventricular dyssynchrony and predicts response to cardiac resynchronization therapy. Heart 2007;93:1034–9.


