**Original Article** | Cardiovascular Imaging

eISSN 2005-8330
https://doi.org/10.3348/kjr.2022.0496
Korean J Radiol 2022;23(12):1251-1259

Korean Journal of Radiology
KJR

# Automated Measurement of Native T1 and Extracellular Volume Fraction in Cardiac Magnetic Resonance Imaging Using a Commercially Available Deep Learning Algorithm

Suyon Chang[1], Kyunghwa Han[2], Suji Lee[2], Young Joong Yang[3], Pan Ki Kim[3], Byoung Wook Choi[2, 3], Young Joo Suh[2]

[1]Department of Radiology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea; [2]Department of Radiology, Research Institute of Radiological Science, Center for Clinical Imaging Data Science, Yonsei University College of Medicine, Seoul, Korea; [3]Phantomics, Inc., Seoul, Korea

**Objective:** T1 mapping provides valuable information regarding cardiomyopathies. Manual drawing is time consuming and prone to subjective errors. Therefore, this study aimed to test a DL algorithm for the automated measurement of native T1 and extracellular volume (ECV) fractions in cardiac magnetic resonance (CMR) imaging with a temporally separated dataset.
**Materials and Methods:** CMR images obtained for 95 participants (mean age ± standard deviation, 54.5 ± 15.2 years), including 36 left ventricular hypertrophy (12 hypertrophic cardiomyopathy, 12 Fabry disease, and 12 amyloidosis), 32 dilated cardiomyopathy, and 27 healthy volunteers, were included. A commercial deep learning (DL) algorithm based on 2D U-net (Myomics-T1 software, version 1.0.0) was used for the automated analysis of T1 maps. Four radiologists, as study readers, performed manual analysis. The reference standard was the consensus result of the manual analysis by two additional expert readers. The segmentation performance of the DL algorithm and the correlation and agreement between the automated measurement and the reference standard were assessed. Interobserver agreement among the four radiologists was analyzed.
**Results:** DL successfully segmented the myocardium in 99.3% of slices in the native T1 map and 89.8% of slices in the post-T1 map with Dice similarity coefficients of 0.86 ± 0.05 and 0.74 ± 0.17, respectively. Native T1 and ECV showed strong correlation and agreement between DL and the reference: for T1, $r$ = 0.967 (95% confidence interval [CI], 0.951–0.978) and bias of 9.5 msec (95% limits of agreement [LOA], -23.6–42.6 msec); for ECV, $r$ = 0.987 (95% CI, 0.980–0.991) and bias of 0.7% (95% LOA, -2.8%–4.2%) on per-subject basis. Agreements between DL and each of the four radiologists were excellent (intraclass correlation coefficient [ICC] of 0.98–0.99 for both native T1 and ECV), comparable to the pairwise agreement between the radiologists (ICC of 0.97–1.00 and 0.99–1.00 for native T1 and ECV, respectively).
**Conclusion:** The DL algorithm allowed automated T1 and ECV measurements comparable to those of radiologists.
**Keywords:** *Magnetic resonance imaging; Heart; Deep learning; T1 mapping; Extracellular volume fraction*

## INTRODUCTION

Cardiac magnetic resonance (CMR) imaging enables the comprehensive evaluation of cardiac structure, function, and tissue characteristics [1]. Late gadolinium enhancement (LGE) is the reference standard for noninvasive imaging of myocardial scars or focal fibrosis in ischemic and non-ischemic cardiomyopathies; however, diffuse myocardial fibrosis can remain undetected because LGE imaging shows the relative difference between enhancing areas and normal nulled myocardium [2-5]. In contrast, T1-mapping provides pixel-wise illustrations of absolute T1 relaxation times on a map and allows direct T1 quantification; therefore, it can detect diffuse myocardial fibrosis that is not assessable by LGE [5]. Native T1 values are a composite signal of myocytes and extracellular volume (ECV). It is increased

**Received:** July 20, 2022 **Revised:** October 5, 2022
**Accepted:** October 6, 2022
**Corresponding author:** Young Joo Suh, MD, PhD, Department of Radiology, Research Institute of Radiological Science, Center for Clinical Imaging Data Science, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.
• E-mail: rongzusuh@gmail.com

in cases of edema, fibrosis, or amyloid deposition, and decreased in cases of lipid or iron overload. When a contrast agent is used, the ECV fraction can be estimated from the ratio of the T1 signal values. It is thought to be a more robust measure of myocardial fibrosis than native T1 or post-contrast T1 [6]. Both native T1 values and ECV provide valuable information for the differential diagnosis and prognosis of various cardiomyopathies [5].

The endocardial and epicardial contours of the left ventricle (LV) must be manually drawn on maps for mapping values. This time-consuming step requires an experienced reviewer to achieve reproducible measurements and to avoid erroneous inclusion of the blood pool or epicardial fat. In addition, manual segmentation is a source of interreader variability. Therefore, an automated method for measuring mapping values is desirable. Automated measurements of native T1 maps have shown good correlation and agreement with manual measurement [7-10]. However, previous studies have mainly focused on the native T1 value, while ECV has rarely been assessed [11].

A recently developed, commercially available deep learning (DL) algorithm allows automated segmentation of T1 maps and measurement of native T1, post-T1, and ECV fractions of the LV myocardium in 16 American Heart Association (AHA) segment models.

In this study, we aimed to test the performance of a commercial DL algorithm for the automated segmentation and measurement of native T1 values and ECV fractions of the myocardium in CMR with a temporally separated dataset and compare its performance with that of radiologists.

## MATERIALS AND METHODS

### Study Design and Participants

This retrospective study was approved by the Institutional Review Board, and the requirement for informed consent was waived (IRB No. 1-2021-0008).

The medical records of 755 eligible patients aged > 19 years who underwent CMR in a tertiary care hospital were retrospectively reviewed from November 2018 to December 2019. Among them, patients with both clinical and CMR-based diagnoses of left ventricular hypertrophy (LVH) (hypertrophic cardiomyopathy [HCM], cardiac amyloidosis [CA], and Fabry disease [FD]) and dilated cardiomyopathy (DCM) were consecutively selected from November 2018. This testing cohort was a temporally separated dataset that did not overlap with the dataset used for DL development and validation. The exclusion criteria were inadequate image quality or uncertain diagnosis (n = 9). Data from a prospective study was used to establish a group of healthy volunteers (HV) [12]. Further details regarding the inclusion criteria are described in the Supplementary Materials section. Finally, 36 patients with LVH [12 with HCM, 12 with FD, and 12 with CA], 32 with DCM, and 27 with HV were included (Fig. 1).

### CMR Acquisition Protocols

CMR imaging was performed using a 3Tesla system

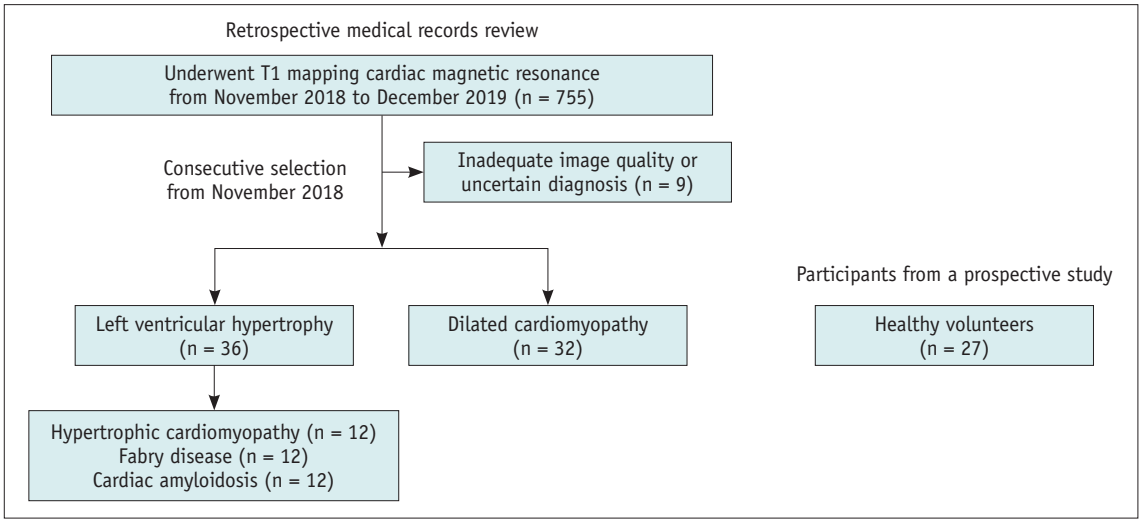


**Fig. 1. Study flowchart.** From 755 patients who underwent cardiac magnetic resonance imaging patients with left ventricular hypertrophy and dilated cardiomyopathy were consecutively selected. The exclusion criteria were inadequate image quality or uncertain diagnosis. Healthy volunteers were included from a prospective study. Finally, 36 patients with left ventricular hypertrophy, 32 patients with dilated cardiomyopathy, and 27 healthy volunteers were included.


https://doi.org/10.3348/kjr.2022.0496

Korean Journal of Radiology
KJR

(Siemens 3T Prismafit). Three short-axis modified look-locker inversion-recovery (MOLLI) images at the basal, mid-ventricular, and apical slices were acquired for native T1 mapping using an 8 image, 11 heart-beat 5-(3)-3 bSSFP sequence with a slice thickness of 8 mm. Next, a total dose of 0.1 mmol/kg gadolinium agent (Uniray, gadoterate meglumine, Dongkook Pharmaceutical Co., Ltd.) was injected. Ten minutes after contrast injection, post-contrast MOLLI T1 mapping was performed for T1 determination at three slices at an identical location as that for native T1 mapping, using a 9 image, 11 heart-beat 4-(1)-3-(1)-2 bSSFP sequence. Motion correction was applied to the native T1 and post-contrast T1 mapping images.

### DL Algorithm for the Automated Measurement of T1 Values and ECV

A DL algorithm based on 2D U-Net (Myomics-T1 software, version 1.0.0, Phantomics) was used for automated analysis of T1 maps. The details of the DL models are described in Supplementary Materials and Figure 2. It automatically segmented the LV myocardium on the maps and provided values of native T1, post-T1, and ECV fractions in 16 AHA segments. To include only the myocardial tissue, a 10% epicardial and endocardial offset was applied. The hematocrit level was entered directly by the user.

### Reference Standard

The LV myocardium was manually segmented by two experienced board-certified cardiac radiologists (with 12 and 25 years of experience, respectively), and the consensus results were used as the reference standard. For manual segmentation, the endocardial and epicardial borders of the LV were delineated using CVI42 software (Circle CVI). The segmentation mask was saved, and T1 times in 16 segments were recorded. Epicardial and endocardial offsets of 10% were applied. The ECV fraction was calculated using the following equation: ECV (%) = $(\Delta R1_m/\Delta R1_b)\cdot(1-Hct)\cdot 100$, where $R1_m$ is R1 in the myocardium, $R1_b$ is R1 in the blood, Hct is the hematocrit level, and ΔR1 is the change in relaxivity. ΔR1 was determined using the following equation: $\Delta R1 = R1_{post}-R1_{pre}$, where $R1_{post}$ and $R1_{pre}$ are R1 values before and after gadolinium chelate administration, respectively [13].

### Segmentation Performance Evaluation

The segmentation performance was evaluated in two ways. First, a cardiac radiologist (with 10 years of experience) reviewed the predicted masks and determined whether segmentation was successful or suboptimal. Image segmentation was considered suboptimal if a myocardium with a valid shape was not produced or if areas other than the LV myocardium were included. Second, we used the Dice similarity coefficient (DSC) to measure the degree of overlap between automated segmentation and the reference standard [14].

### Reader Study

A reader study was performed with four radiologist readers, including two radiology residents (R1 and R2 [non-authors] with 3 and 4 years of experience, respectively) and two board-certified cardiac radiologists (R3 and R4 with 6 and 10 years of experience, respectively). They independently segmented the LV myocardium of 95 subjects using the CVI42 software (Circle CVI). They were blinded to the reference standards or results of other readers. T1 times in the 16 AHA segments were recorded, and the ECV fraction

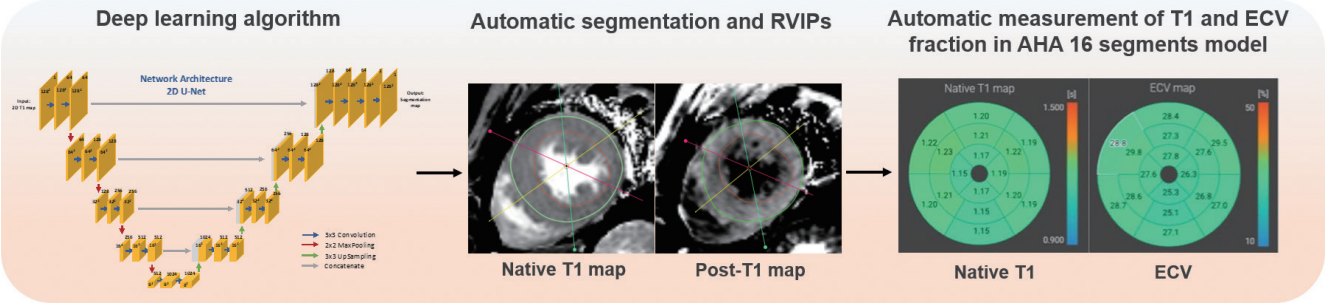


**Fig. 2. Illustration of the fully automated myocardium analysis.** The DL architecture, 2D U-Net, used for myocardium segmentation is shown in the left image. Developed DL models segment the myocardium in native T1 and post-T1 maps. Based on DL segmentation results, RVIPs and reference axes are generated for AHA 16 segments analysis. In the middle image, automatically defined myocardium contours and color-coded reference lines are shown. T1 maps and ECV fractions are converted to the AHA 16 segments model. Automated measurements are summarized in the form of a bull's eye map as shown in the right image. AHA = American Heart Association, DL = deep learning, ECV = extracellular volume, RVIP = right ventricular insertion point


https://doi.org/10.3348/kjr.2022.0496

was calculated as described above.

### Statistical Analysis

Segmentation accuracy of the DL algorithm in terms of DSC was compared among patients with left ventricular hypertrophy, dilated cardiomyopathy, and HV using the Kruskal–Wallis test and post-hoc tests with Bonferroni correction. Correlations between the automated and reference values were assessed using scatter plots with regression slopes and Pearson's correlation coefficients. Agreements between automated and reference values were evaluated using Bland–Altman analyses. Interobserver agreement in the reader study was assessed using the intraclass correlation coefficient (ICC) with a two-way random model under absolute agreement. Statistical analyses were performed using the R software (version 4.0.5. (R Foundation for Statistical Computing). Statistical significance was set at $p < 0.05$.

## RESULTS

### Baseline Characteristics

Table 1 presents the baseline characteristics of the study participants. The HVs were the youngest among the three groups (47.7 vs. 56.3–57.9 years, $p = 0.019$). The height, body mass index, body surface area, and hematocrit were not significantly different between the groups. Patients with DCM had the highest body weight, but the difference was clinically insignificant (mean, 68.8 vs. 61.9–62.4 kg, $p = 0.043$).

### Segmentation Performance of DL Algorithm

The computation time for segmenting each map was less than 5 seconds (mean 1.4 seconds, range 0.6–4.6 seconds in the native T1 map; mean 0.6 seconds, range 0.6–1.1 seconds in the post-T1 map). DL successfully segmented the myocardium in the native T1 map in 99.3% (283/285) of the slices in 97.9% (93/95) of patients. Suboptimal segmentations were noted in one patient with CA (basal slice) and one patient with DCM (apical slice). In the post-T1 map, DL successfully segmented the myocardium in 89.8% (256/285) of the slices in 87.4% (83/95) of subjects. Suboptimal segmentations were present in 9 of 12 patients with CA (all slices in 8 patients, mid and apical slices in 1 patient), 1 patient with FD (apical slice), 1 patient with DCM (apical slice), and 1 patient with HV (apical slice).

The DSC was 0.86 ± 0.05 in the native T1 map and 0.74 ± 0.17 in the post-T1 map (Table 2, Supplementary Fig. 1). The DSC on the native T1 map was significantly different among the three groups (mean 0.88, 0.85, and 0.85 for LVH, DCM, and HV, respectively; $p < 0.001$). In post-hoc analysis, DSC was higher in the LVH group than in the DCM or HV groups (adjusted $p = 0.001$ for both). The DSC on the post-T1 map was not significantly different among the three groups (mean 0.68, 0.76, and 0.80 for LVH, DCM, and HV, respectively; $p = 0.134$). In the LVH group, the mean DSC on the post-T1 map was 0.81, 0.84, and 0.40 in HCM, FD, and CA, respectively. If the results of patients with CA were excluded from the analysis, DSC on the post-T1 map was 0.79 ± 0.08 in the total population (n = 83) and 0.83 ± 0.05

**Table 1. Baseline Characteristics of Study Participants**

| Parameter | All Subjects (n = 95) | Left Ventricular Hypertrophy (n = 36) | Dilated Cardiomyopathy (n = 32) | Healthy Volunteers (n = 27) | P* |
|---|---|---|---|---|---|
| Sex | | | | | 0.507 |
| Female | 48 (50.5) | 16 (44.4) | 16 (50) | 16 (59.3) | |
| Male | 47 (49.5) | 20 (55.6) | 16 (50) | 11 (40.7) | |
| Age, years | 54.5 ± 15.2 | 57.9 ± 12.9 | 56.3 ± 15.8 | 47.7 ± 15.7 | 0.019 |
| Height, cm | 164.6 ± 8.6 | 163.2 ± 6.9 | 166.5 ± 10.6 | 164.2 ± 7.9 | 0.286 |
| Weight, kg | 64.4 ± 12.2 | 62.4 ± 9.0 | 68.8 ± 15.3 | 61.9 ± 10.7 | 0.043 |
| BMI, kg/m$^2$ | 23.7 ± 3.6 | 23.4 ± 3.2 | 24.7 ± 4.4 | 22.8 ± 2.6 | 0.106 |
| BSA, m$^2$ | 1.7 ± 0.2 | 1.7 ± 0.1 | 1.8 ± 0.2 | 1.7 ± 0.2 | 0.060 |
| Hematocrit, % | 39.8 ± 5.3 | 38.5 ± 6.0 | 40.4 ± 5.2 | 40.9 ± 4.1 | 0.154 |
| LVEDVi, mL/m$^2$ | 100.4 ± 44.1 | 79.6 ± 22.7 | 150.5 ± 35.3 | 68.7 ± 11.3 | < 0.001 |
| LVESVi, mL/m$^2$ | 58.6 ± 46.2 | 32.0 ± 14.4 | 115.5 ± 33.8 | 26.8 ± 6.1 | < 0.001 |
| LVEF, % | 48.1 ± 19.5 | 59.9 ± 11.7 | 23.9 ± 7.7 | 61.2 ± 5.4 | < 0.001 |

Data are mean ± standard deviation or number (%). *Comparison among three groups. BMI = body mass index, BSA = body surface area, LVEDVi = left ventricular end-diastolic volume index, LVEF = left ventricular ejection fraction, LVESVi = left ventricular end-systolic volume index


https://doi.org/10.3348/kjr.2022.0496

Korean Journal of Radiology
KJR

in the LVH group (n = 24).

### Correlation and Agreement between Automated and Reference Values

There was a strong correlation between the automated and reference native T1 values per subject ($r$ = 0.967 [95% confidence interval {CI} 0.951–0.978], slope = 0.9338), per-slice ($r$ = 0.980 [95% CI 0.975–0.984], slope = 0.9957), and per-segment analyses ($r$ = 0.951 [95% CI 0.946–0.956]; slope = 0.9648) (Fig. 3, top row).

The automated and reference native T1 values were in good agreement in per-subject (bias 9.5 msec, 95% limits of agreement [LoA] -23.6–42.6 msec), per-slice (bias 10.6 msec, 95% LoA -32.8–53.9 msec), and per-segment analyses (bias 9.6 msec, 95% LoA -61.4–80.6 msec) (Fig. 3, bottom row).

Automated ECV was strongly correlated with the reference ECV per subject ($r$ = 0.987 [95% CI 0.980–0.991],

**Table 2. Segmentation Performances of the Deep Learning Algorithm on Native T1 and Post-T1 Maps**

| Group | DSC on Native T1 Map | DSC on Post-T1 Map |
|---|---|---|
| All subjects (n = 95) | 0.86 ± 0.05 | 0.74 ± 0.17 |
| Left ventricular hypertrophy (n = 36) | 0.88 ± 0.05 | 0.68 ± 0.25 |
| Hypertrophic cardiomyopathy (n = 12) | 0.89 ± 0.02 | 0.81 ± 0.05 |
| Fabry disease (n = 12) | 0.87 ± 0.08 | 0.84 ± 0.04 |
| Cardiac amyloidosis (n = 12) | 0.87 ± 0.04 | 0.40 ± 0.24 |
| Dilated cardiomyopathy (n = 32) | 0.85 ± 0.05 | 0.76 ± 0.10 |
| Healthy volunteers (n = 27) | 0.85 ± 0.03 | 0.80 ± 0.07 |
| $p$ value* | < 0.001 | 0.134 |
| Adjusted $p$ value† | 0.001 | NA |
| Adjusted $p$ value‡ | 0.001 | NA |

Data are mean ± standard deviation. *Comparison among left ventricular hypertrophy, dilated cardiomyopathy, and healthy volunteers, †Comparison between left ventricular hypertrophy and dilated cardiomyopathy, ‡Comparison between left ventricular hypertrophy and healthy volunteers. DSC = Dice similarity coefficient, NA = not applicable

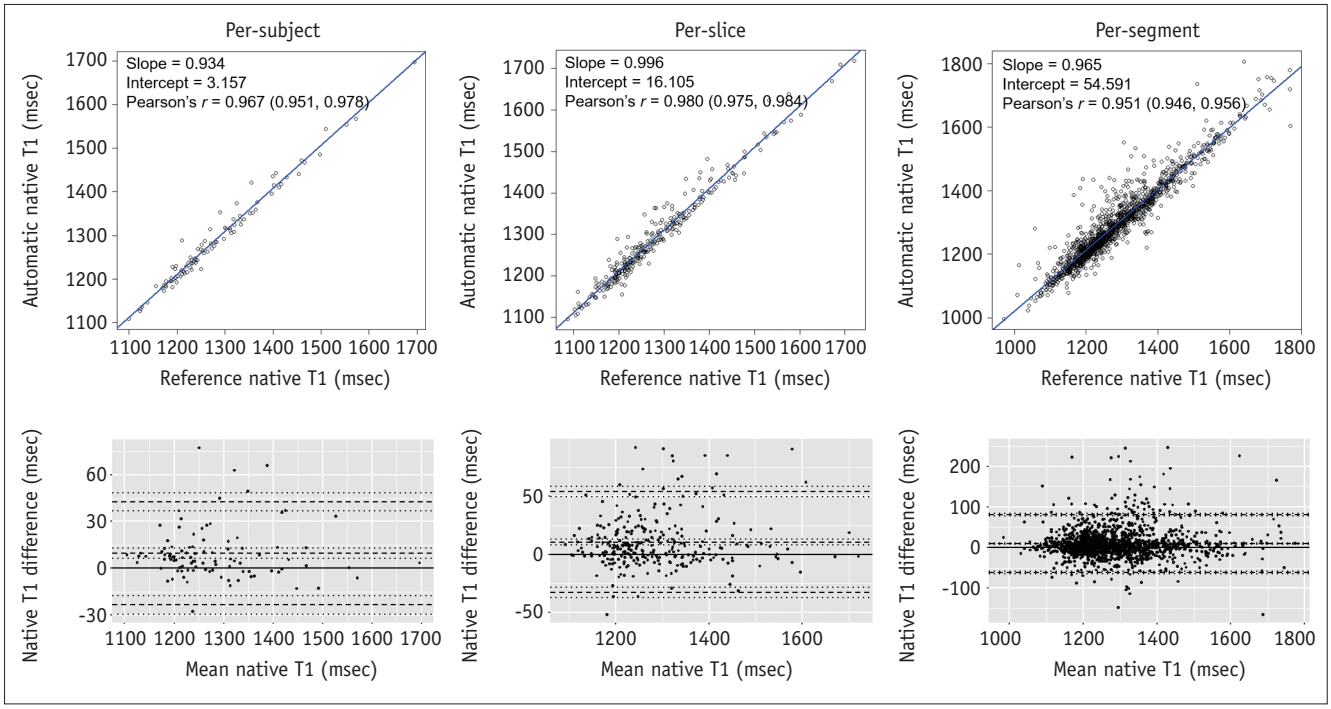


**Fig. 3. Scatter plots and Bland-Altman plots of automated and reference native T1 values.** Automated T1 showed strong correlation with reference T1 in per-subject ($r$ = 0.967 [95% CI 0.951–0.978], slope = 0.934), per-slice ($r$ = 0.980 [95% CI 0.975–0.984], slope = 0.996), and per-segment analyses ($r$ = 0.951 [95% CI 0.946–0.956], slope = 0.965). Automated and reference T1 values were in good agreement in per-subject (bias 9.5 msec, 95% LoA -23.6–42.6 msec), per-slice (bias 10.6 msec, 95% LoA -32.8–53.9 msec), and per-segment analyses (bias 9.6 msec, 95% LoA -61.4–80.6 msec). CI = confidence interval, LoA = limits of agreement


https://doi.org/10.3348/kjr.2022.0496

slope = 0.943), per-slice ($r$ = 0.968 [95% CI 0.960–0.975], slope = 0.943), and per-segment analyses ($r$ = 0.956 [95% CI 0.951–0.960], slope = 0.935) (Fig. 4, top row).

The automated and reference ECV were in good agreement in per-subject (bias 0.7%, 95% LoA -2.8%–4.2%), per-slice (bias 0.9%, 95% LoA -4.4%–6.2%), and per-segment analyses (bias 0.7%, 95% LoA -5.7%–7.2%) (Fig. 4, bottom row).

### Interobserver Agreement and Agreement between the Automated Measurement and Radiologist Readers

Regarding the native T1 value, pairwise agreement between radiologists was excellent (ICC range, 0.97–1.00) (Table 3), and the ICC among the four readers was 0.99. The agreement between the automated measurement and each radiologist was excellent, with an ICC range 0.98–0.99.

Regarding ECV, the pairwise agreement between radiologists was excellent (ICC range, 0.99–1.00), and the ICC among the four readers was 0.99. The agreement between the DL and each radiologist was excellent, with an ICC range 0.98–0.99.

Supplementary Table 1 shows the inter-observer agreement among the three groups. In the LVH group, agreement was excellent in all pairwise comparisons (ICC 0.99–1.00). In the DCM and HV groups, the agreement between DL and each radiologist was good to excellent for native T1 (ICC 0.83–0.95) and ECV (ICC 0.78–0.97), while the pairwise agreement between radiologists was moderate

**Table 3. Interobserver Agreement and Agreement between the Automated and Manually Measured Native T1 and ECV Fraction**

| Comparison | ICC (95% Confidence Interval)* | |
|---|---|---|
| | Native T1 | ECV |
| Reader 1–reader 2 | 0.98 (0.78–1.00) | 0.99 (0.86–1.00) |
| Reader 1–reader 3 | 0.97 (0.59–0.99) | 0.99 (0.88–1.00) |
| Reader 1–reader 4 | 1.00 (0.97–1.00) | 1.00 (1.00–1.00) |
| Reader 2–reader 3 | 1.00 (0.97–1.00) | 1.00 (1.00–1.00) |
| Reader 2–reader 4 | 1.00 (1.00–1.00) | 1.00 (0.99–1.00) |
| Reader 3–reader 4 | 0.99 (0.97–1.00) | 1.00 (0.99–1.00) |
| Among four readers | 0.99 (0.96–0.99) | 0.99 (0.99–1.00) |
| Automatic–reader 1 | 0.98 (0.95–0.99) | 0.99 (0.98–0.99) |
| Automatic–reader 2 | 0.99 (0.98–0.99) | 0.98 (0.97–0.99) |
| Automatic–reader 3 | 0.98 (0.95–0.99) | 0.99 (0.97–0.99) |
| Automatic–reader 4 | 0.99 (0.98–0.99) | 0.99 (0.98–0.99) |
| Automatic–average of four readers | 0.99 (0.99–0.99) | 0.99 (0.98–0.99) |

*Based on per-subject basis. ECV = extracellular volume, ICC = intraclass correlation coefficient

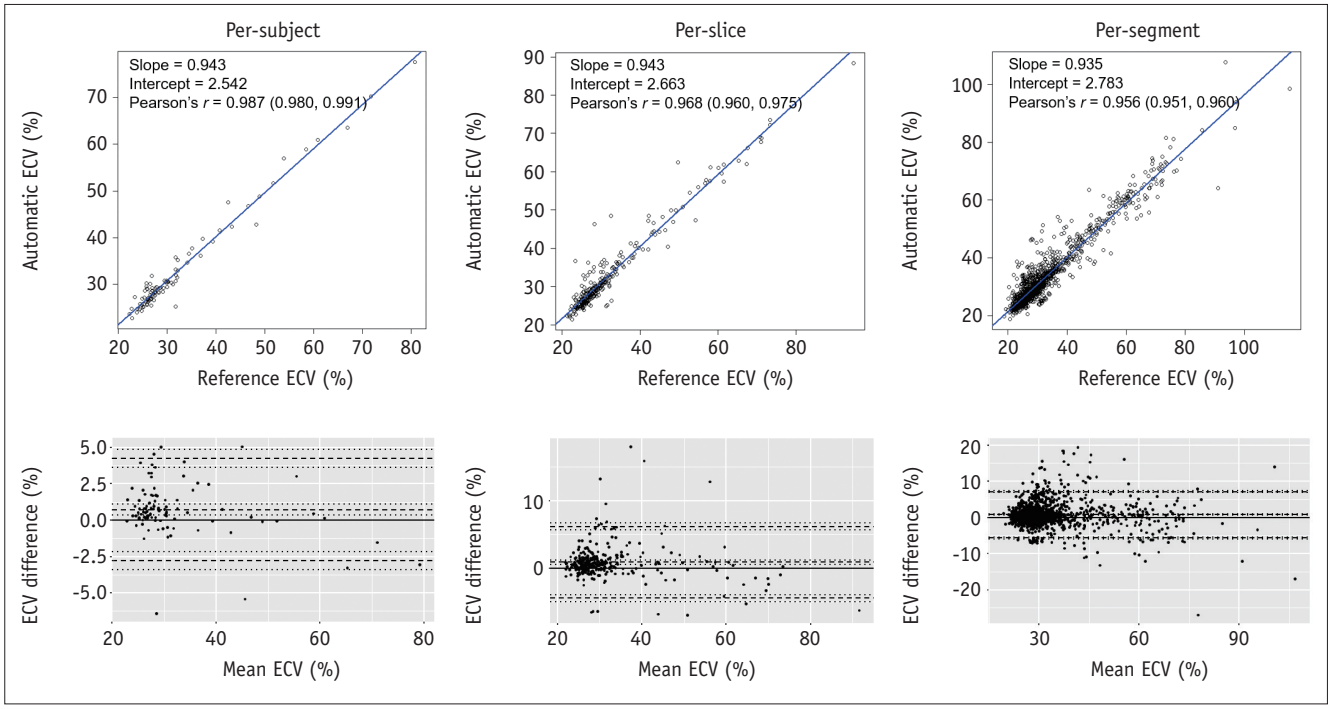


**Fig. 4. Scatter plots and Bland-Altman plots of automated and reference ECV.** Automated ECV showed strong correlation with reference ECV in per-subject ($r$ = 0.987 [95% CI 0.980–0.991], slope = 0.943), per-slice ($r$ = 0.968 [95% CI 0.960–0.975], slope = 0.943), and per-segment analyses ($r$ = 0.956 [95% CI 0.951–0.960], slope = 0.935). Automated and reference ECV were in good agreement in per-subject (bias 0.7%, 95% LoA -2.8%–4.2%), per-slice (bias 0.9%, 95% LoA -4.4%–6.2%), and per-segment analyses (bias 0.7%, 95% LoA -5.7%–7.2%). CI = confidence interval, ECV = extracellular volume, LoA = limits of agreement


https://doi.org/10.3348/kjr.2022.0496

Korean Journal of Radiology
KJR

to excellent for native T1 (ICC 0.61–0.99) and ECV (ICC 0.66–0.99).

## DISCUSSION

In this study, we aimed to test a DL algorithm for the automated measurement of myocardial native T1 values and ECV fractions in CMR with a temporally separated dataset. The algorithm achieved fast segmentation (< 5 sec/image) and successfully segmented the myocardium in 99.3% of slices in the native T1 map and 89.8% of slices in the post-T1 map with high DSC (0.86 ± 0.05 and 0.74 ± 0.17, respectively). Native T1 and ECV from the automated method and the reference standard were highly correlated with good agreement. Further, DL achieved very good agreement with the radiologists, similar to the inter-radiologist agreement.

Automatic segmentation based on DL reduces the workload of radiologists, minimizes user bias, and enables large-scale study. Although extensive research has been conducted on the automated segmentation of cine images [15-17], information on the automated analysis of T1 mapping CMR is lacking [7-11,18]. In 2018, Huang et al. [18] proposed an automated regional analysis of native T1 values in the LV. However, only ten healthy individuals were included in this study. Prior studies have mainly focused on the native T1 map [7,8,10,18], while only a few studies have investigated the post-T1 map [9] or ECV [11]. The performance of DL has been evaluated in a heterogeneous population [7,8,11] or in patients with known myocardial fibrosis who have LGE [9]. In addition, most studies did not include a per-segment analysis [7-10,18], which may be critical for analyzing focal myocardial disease. Most importantly, these studies are still in the research stage and have not yet been used widely in clinical practice. We comprehensively evaluated automated native T1 and ECV measurements in LVH, DCM, and HV on a per-subject, per-slice, and per-segment basis using a commercially available DL algorithm with experts' measurements as reference.

DL successfully segmented the myocardium in most slices (99.3%) in the native T1 map. This success rate was higher than that reported in previous studies on the automated segmentation of native T1 maps (84.3%–91.3%) [7,8]. In the post-T1 map, DL successfully segmented the myocardium in 89.9% of slices. The most suboptimal segmentation occurred in patients with CA or apical slices. The DSC on the native T1 map in our study (0.86 ± 0.05) was similar to the previously reported values of automated measurement (range, 0.81–0.85) [7-9]. On the post-T1 map, the mean DSC was 0.74 ± 0.17, similar to the value reported in a previous study that included patients with ischemic cardiomyopathy, DCM, and HCM (range, 0.74–0.77) [9]. If patients with CA were excluded from our analysis, a higher DSC (0.79 ± 0.08) was obtained.

Our results suggest that automated segmentation is challenging in the apical slices and post-T1 maps of patients with CA. Difficulty in automated segmentation of the apical myocardium has also been reported in a previous study [7]. This is probably because the apical slice is more susceptible to partial volume effects [19], which requires cautious segmentation. Using a cross-reference with a long-axis image could be helpful in this regard. If the location of the apical slice identified in the long-axis image is inappropriate, or there is a high possibility that a partial volume effect will occur, the slice may be excluded from the analysis. Additionally, if the T1 map is obtained from four or more slices, appropriate slices can be selected from the long-axis image. Myocardial segmentation on the post-T1 map can be more challenging than on the native T1 map, mainly because of the low contrast between the myocardium and the blood pool. Furthermore, in cases of amyloid deposition or myocardial fibrosis, gadolinium accumulates in the expanded extracellular space, resulting in shortening of the T1 inversion time [20,21]. This further complicates segmentation of the post-T1 map. In fact, the DSC on the post-T1 map was the lowest in the LVH group among the three groups, most likely because of the low DSC in patients with CA. Therefore, the automated segmentation of the post-T1 map has limitations in patients with CA, and a re-check of the segmentation is necessary in such cases.

Automated measurements of T1 values and ECV fractions revealed a strong correlation and good agreement with the reference standard. However, some cases showed significant differences between the two methods. While previous studies excluded cases with unsuccessful segmentation for correlation or reproducibility analysis [7,8,11], we analyzed all cases, including cases with suboptimal segmentation. This resulted in some outliers in the Bland–Altman plots.

Automated measurements of native T1 and ECV also achieved excellent agreement with the radiologists in pairwise comparisons, showing rates similar to the inter-radiologist agreement. In addition, it consistently achieved good-to-excellent agreement with the radiologists' values in the subgroup analysis of all three groups. Our results

 https://doi.org/10.3348/kjr.2022.0496 

suggest that this DL algorithm may reach the radiologist's level of agreement for measuring mapping values. Therefore, it may reduce the workload of radiologists in clinical practice and can be beneficial for research that requires analyzing a large amount of data with reduced reader dependency.

Our study had several limitations. First, this was a retrospective study that tested a commercial algorithm. Second, all images were acquired using a single MR scanner at 3T. Further studies are required to evaluate the performance of images acquired at different field strengths and from different machines. Third, this study included only cases without significant artifacts, and the performance of this DL algorithm may differ for patients with severe artifacts. Finally, because this DL algorithm does not have an automated quality control process, an inadequate segmentation can be missed. However, an overall good correlation and agreement with the reference were achieved, and agreements between the DL and each radiologist were comparable to the inter-radiologist agreements. A quick review of automatically drawn regions of interest is desirable in clinical practice, and modifications may be needed in some cases, especially in patients with suspected CA. Nevertheless, the overall burden of work can be reduced compared with manual processing.

In conclusion, automated T1 and ECV measurements using the commercial DL algorithm showed good correlation and agreement with the reference standard and radiologist readers.

## Supplement

The Supplement is available with this article at https://doi.org/10.3348/kjr.2022.0496.

### Availability of Data and Material

The datasets generated or analyzed during the study are available from the corresponding author on reasonable request.

### Conflicts of Interest

Kyunghwa Han who is on the editorial board of the *Korean Journal of Radiology* was not involved in the editorial evaluation or decision to publish this article.
Pan Ki Kim and Byoung Wook Choi are founders of Phantomics, Inc. (Seoul, Korea) and Young Joong Yang is an employee of the same company. The company supported the software for this study. Other authors have no conflicts of interest to declare.

### Author Contributions

Conceptualization: Suyon Chang, Kyunghwa Han, Young Joo Suh. Data curation: Suyon Chang, Young Joo Suh. Formal analysis: Suyon Chang, Kyunghwa Han. Funding acquisition: Young Joo Suh. Investigation: Suyon Chang, Kyunghwa Han, Suji Lee, Young Joo Suh. Methodology: Suyon Chang, Kyunghwa Han, Young Joo Suh. Project administration: Young Joo Suh. Resources: Young Joo Suh. Software: Young Joong Yang, Pan Ki Kim, Byoung Wook Choi. Supervision: Young Joo Suh. Validation: Suyon Chang, Kyunghwa Han, Young Joo Suh. Visualization: Suyon Chang. Writing—original draft: Suyon Chang. Writing—review & editing: all authors.

### ORCID iDs

Suyon Chang
https://orcid.org/0000-0002-9221-8116
Kyunghwa Han
https://orcid.org/0000-0002-5687-7237
Suji Lee
https://orcid.org/0000-0002-8770-622X
Young Joong Yang
https://orcid.org/0000-0003-2231-066X
Pan Ki Kim
https://orcid.org/0000-0001-6006-0732
Byoung Wook Choi
https://orcid.org/0000-0002-8873-5444
Young Joo Suh
https://orcid.org/0000-0002-2078-5832

### Funding Statement

This research was supported by the Technology development Program (S3033533) funded by the Ministry of SMEs and Startups (MSS, Korea).

## REFERENCES

1. Captur G, Manisty C, Moon JC. Cardiac MRI evaluation of myocardial disease. *Heart* 2016;102:1429-1435
2. Kim RJ, Wu E, Rafael A, Chen EL, Parker MA, Simonetti O, et al. The use of contrast-enhanced magnetic resonance imaging to identify reversible myocardial dysfunction. *N Engl J Med* 2000;343:1445-1453
3. Ibanez B, Aletras AH, Arai AE, Arheden H, Bax J, Berry C, et al. Cardiac MRI endpoints in myocardial infarction

experimental and clinical trials: JACC scientific expert panel. *J Am Coll Cardiol* 2019;74:238-256
4. Kuruvilla S, Adenaw N, Katwal AB, Lipinski MJ, Kramer CM, Salerno M. Late gadolinium enhancement on cardiac magnetic resonance predicts adverse cardiovascular outcomes in nonischemic cardiomyopathy: a systematic review and meta-analysis. *Circ Cardiovasc Imaging* 2014;7:250-258
5. Haaf P, Garg P, Messroghli DR, Broadbent DA, Greenwood JP, Plein S. Cardiac T1 mapping and extracellular volume (ECV) in clinical practice: a comprehensive review. *J Cardiovasc Magn Reson* 2016;18:89
6. Moon JC, Messroghli DR, Kellman P, Piechnik SK, Robson MD, Ugander M, et al. Myocardial T1 mapping and extracellular volume quantification: a Society for Cardiovascular Magnetic Resonance (SCMR) and CMR working group of the European Society of Cardiology consensus statement. *J Cardiovasc Magn Reson* 2013;15:92
7. Fahmy AS, El-Rewaidy H, Nezafat M, Nakamori S, Nezafat R. Automated analysis of cardiovascular magnetic resonance myocardial native T1 mapping images using fully convolutional neural networks. *J Cardiovasc Magn Reson* 2019;21:7
8. Puyol-Antón E, Ruijsink B, Baumgartner CF, Masci PG, Sinclair M, Konukoglu E, et al. Automated quantification of myocardial tissue characteristics from native T1 mapping using neural networks with uncertainty-based quality-control. *J Cardiovasc Magn Reson* 2020;22:60
9. Farrag NA, Lochbihler A, White JA, Ukwatta E. Evaluation of fully automated myocardial segmentation techniques in native and contrast-enhanced T1-mapping cardiovascular magnetic resonance images using fully convolutional neural networks. *Med Phys* 2021;48:215-226
10. Hann E, Popescu IA, Zhang Q, Gonzales RA, Barutçu A, Neubauer S, et al. Deep neural network ensemble for on-the-fly quality control-driven segmentation of cardiac MRI T1 mapping. *Med Image Anal* 2021;71:102029
11. Zhu Y, Fahmy AS, Duan C, Nakamori S, Nezafat R. Automated myocardial T2 and extracellular volume quantification in cardiac MRI using transfer learning–based myocardium segmentation. *Radiol Artif Intell* 2020;2:e190034
12. Suh YJ, Kim PK, Park J, Park EA, Jung JI, Choi BW. Phantom-based correction for standardization of myocardial native T1 and extracellular volume fraction in healthy subjects at 3-Tesla cardiac magnetic resonance imaging. *Eur Radiol* 2022 Jun 30 [Epub]. https://doi.org/10.1007/s00330-022-08936-8
13. Diesbourg LD, Prato FS, Wisenberg G, Drost DJ, Marshall TP, Carroll SE, et al. Quantification of myocardial blood flow and extracellular volumes using a bolus injection of Gd-DTPA: kinetic modeling in canine ischemic disease. *Magn Reson Med* 1992;23:239-253
14. Park SH, Choi J, Byeon JS. Key principles of clinical validation, device approval, and insurance coverage decisions of artificial intelligence. *Korean J Radiol* 2021;22:442-453
15. Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W, et al. Deep learning for cardiac image segmentation: a review. *Front Cardiovasc Med* 2020;7:25
16. Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med Image Anal* 2017;35:159-171
17. Isensee F, Jaeger PF, Full PM, Wolf I, Engelhardt S, Maier-Hein KH. Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features. Proceedings of the Statistical Atlases and Computational Models of the Heart ACDC and MMWHS Challenges; 2017 Sep 10-14; Quebec, Canada: Springer International Publishing; 2018; p. 120-129
18. Huang HH, Huang CY, Chen CN, Wang YW, Huang TY. Automatic regional analysis of myocardial native T1 values: left ventricle segmentation and AHA parcellations. *Int J Cardiovasc Imaging* 2018;34:131-140
19. Lee HY, Codella N, Cham M, Prince M, Weinsaft J, Wang Y. Left ventricle segmentation using graph searching on intensity and gradient and a priori knowledge (lvGIGA) for short-axis cardiac magnetic resonance imaging. *J Magn Reson Imaging* 2008;28:1393-1401
20. Maceira AM, Joshi J, Prasad SK, Moon JC, Perugini E, Harding I, et al. Cardiovascular magnetic resonance in cardiac amyloidosis. *Circulation* 2005;111:186-193
21. Parsai C, O'Hanlon R, Prasad SK, Mohiaddin RH. Diagnostic and prognostic value of cardiovascular magnetic resonance in non-ischaemic cardiomyopathies. *J Cardiovasc Magn Reson* 2012;14:54