

Supplemental materials

1. Patient characteristics

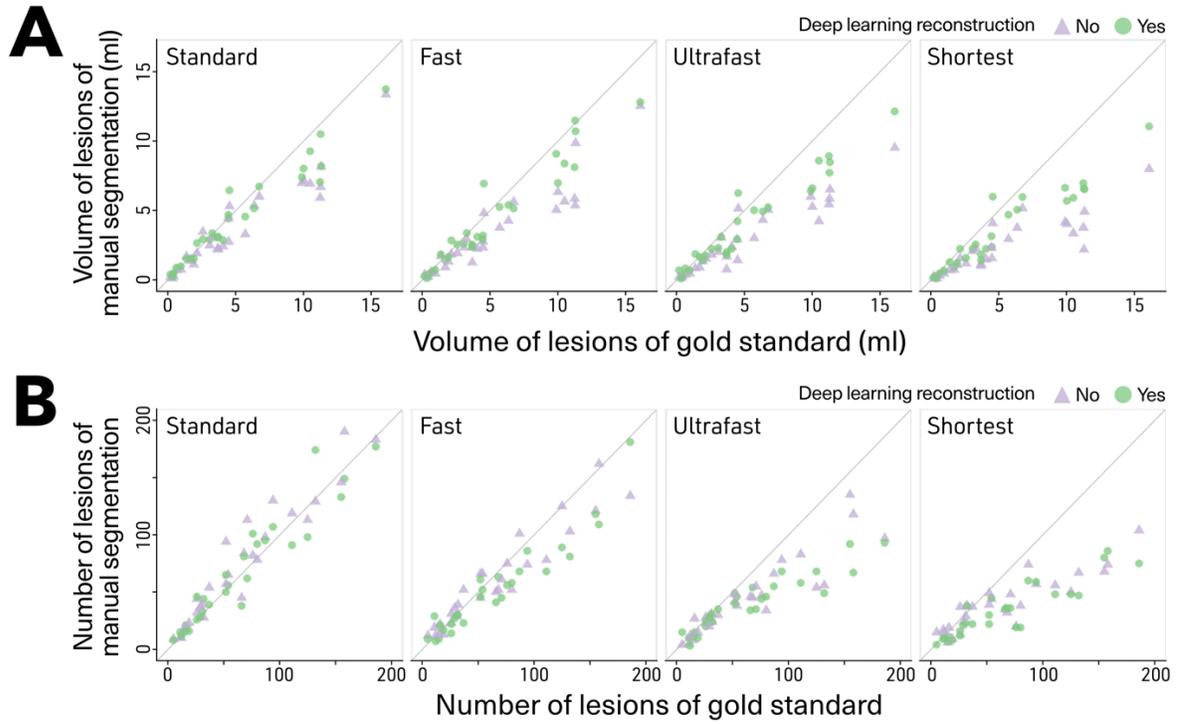
Characteristics	N = 28
Mean age, years (range)	41 (24 – 56)
Gender	
Male	4 (14%)
Female	24 (86%)
Median disease duration, years (IQR)	13 (5 – 17)
Median EDSS score (IQR)	2.0 (1.5 – 3.0)
Number of relapses during the last year	0
Current disease-modifying therapy	
ocrelizumab	1 (3.6%)
natalizumab	27 (96.4%)
Median lesion-load volume, ml (IQR)	3.68 (1.77 – 7.14)

Supplemental Table 1

EDSS: expanded disability status scale; DMT: disease-modifying therapy; IQR: interquartile range.

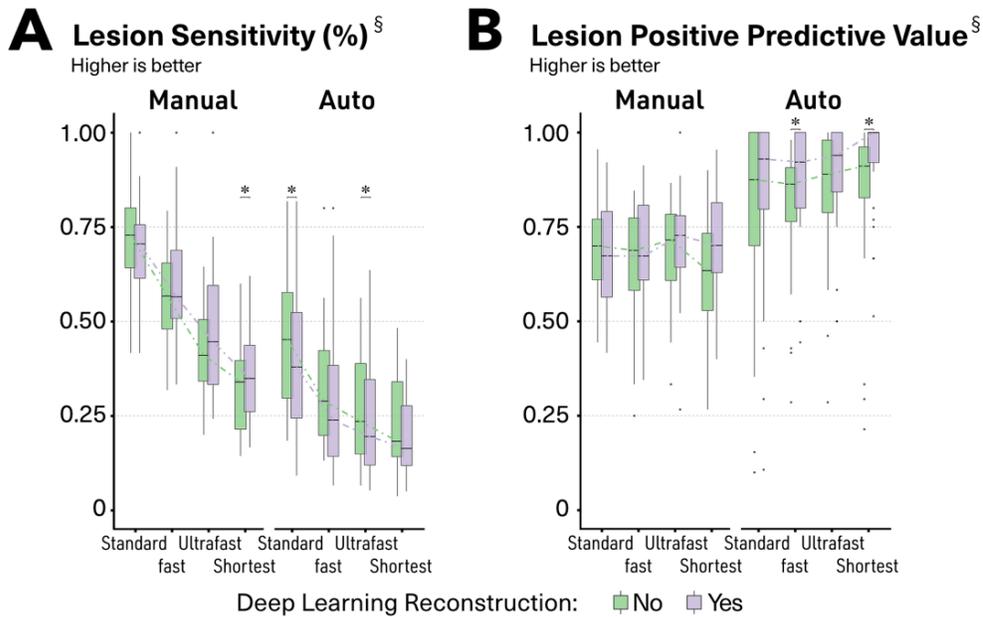
2. Supplemental data on performance metrics

This is another representation of the comparison between the manual segmentation and the gold standard. The number and the volume of lesions were clearly lower than gold standard in faster sequences. This was improved after dDLR especially for the volume of the lesions and to a lesser degree for the number of lesions.



Supplemental Figure 1

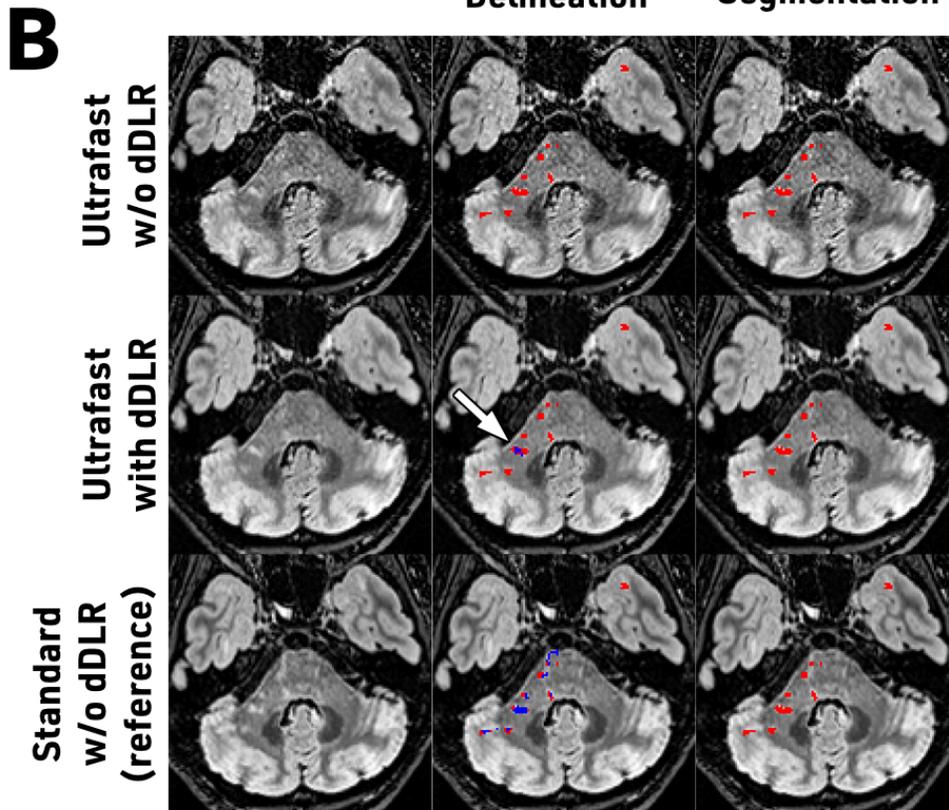
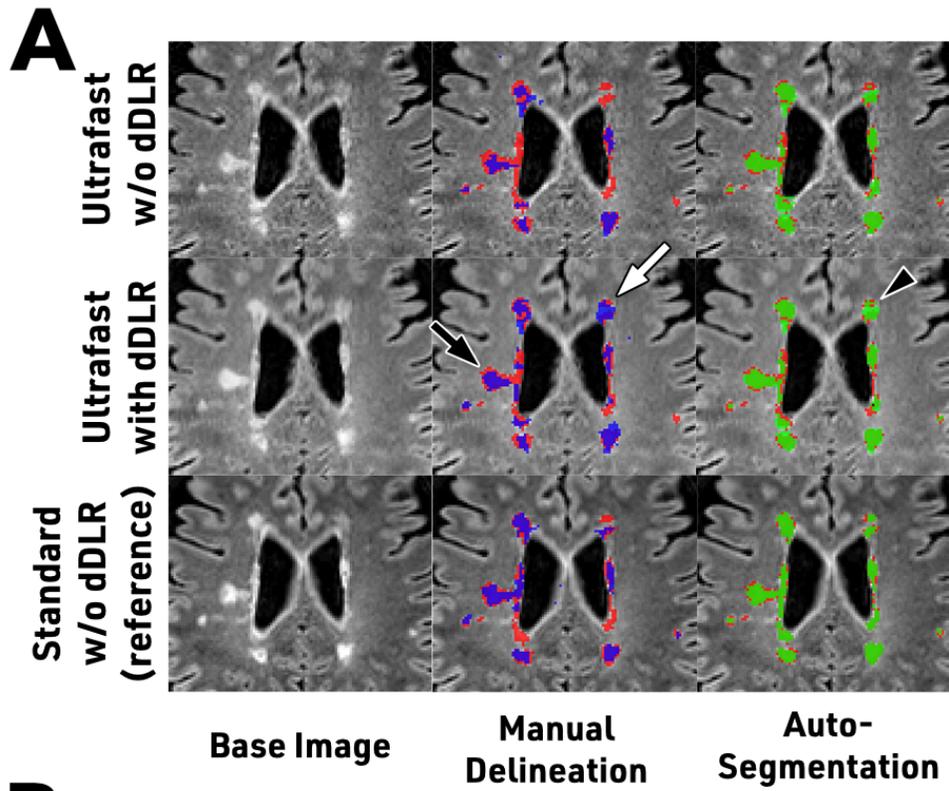
Scatter plot for the number and the volume of lesions. Each point represents a patient.



Supplemental Figure 2

Lesion sensitivity and lesion positive predictive value (lesion-wise metrics).

3. Supplemental illustrations of the effects of dDLR



Supplemental Figure 3

Additional illustrative examples of segmentations for supra-tentorial (A) and posterior fossa lesions (B). Illustrative axial slices of ultrafast FLAIR without and with dDLR. Standard FLAIR without dDLR is also shown for reference. The red mask represents the gold standard that comes from the delineation of standard FLAIR by two expert readers; the blue mask is the manual delineation from a third reader; and the green mask is the automatic segmentation from volBrain software. After applying dDLR, some lesions showed contours closer to the gold standard (black arrows and arrowheads). dDLR also retrieved lesions that were missed on the original image (white arrows).
dDLR: denoising using deep learning-based reconstruction.

4. Evaluation of Apparent Signal-to-Noise Ratio (SNR) and Contrast-to-Noise Ratio (CNR)

4.1 Evaluation of patients' images

We measured the apparent SNR for each sequence using the following definition:

$$\text{apparent SNR} = \frac{\text{mean}(SI_{NWM})}{\text{std}(SI_{BG})}$$

where SI_{NWM} refers to the signal intensity of normal-appearing white matter from 3 regions of interest (ROIs) placed on one slice at the level of the corona radiata, and $\text{std}(SI_{BG})$ refers to the standard deviation of the background noise. The background ROI was set close to the skull because of peripheral signal suppression during the parallel imaging reconstruction process in this vendor's machine. This ROI was placed on top of the head to avoid unfolding miscalculation due to parallel imaging in phase- and slice-encoding directions (Supplemental Figure 1).

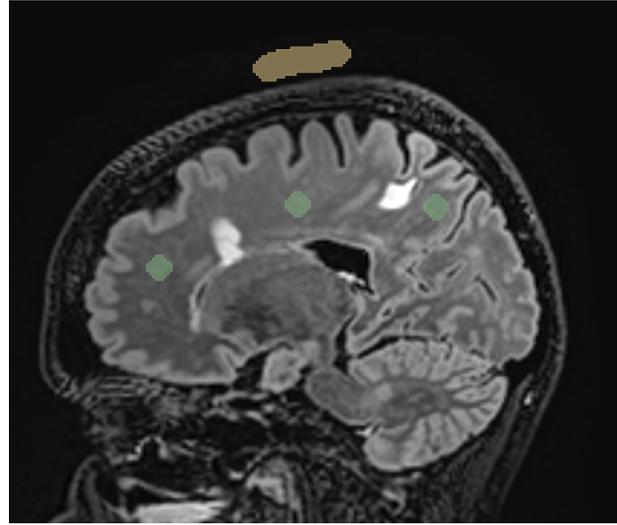
We also measured the lesion CNR for each sequence using the following definition:

$$CNR_{\text{lesion}} = \frac{\text{mean}(SI_{\text{lesion}}) - \text{mean}(SI_{NWM})}{\text{std}(SI_{BG})}$$

where SI_{lesion} refers to signal intensity from the lesions as defined using the gold standard lesion mask on one slice at the level of the corona radiata or at the level of the striatum depending on the lesion load.

Supplemental Figure 4

The green areas are the ROIs of normal-appearing white matter. The yellow area is the ROI of background noise, placed in the readout direction. For the ROI of multiple sclerosis lesions, the mask of ground truth was used (not shown in this figure).



4.2 Evaluation Fruits' Phantom

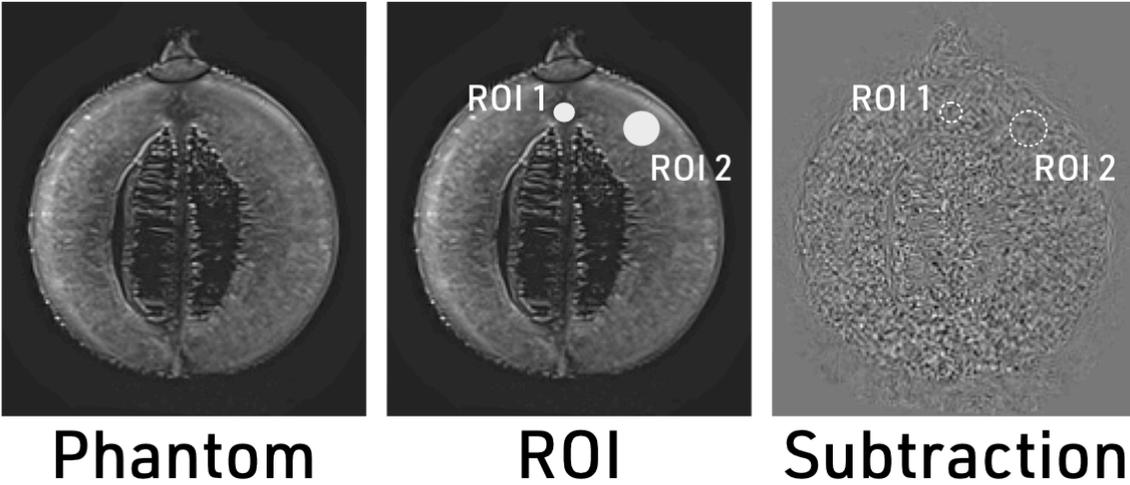
We also performed the experiments using a fruit phantom. The phantom study allows multiple scans to quantify the noise more accurately than what can be done with ROIs in the air in vivo that can be corrupted when a parallel imaging technique is used. All four 3D-FLAIR sequences (standard, fast, ultrafast, and shortest) were scanned twice each, and dDLR was applied to each series. These procedures were repeated five times, and apparent SNR and CNR were measured by averaging the results of the last four sets. Signals were evaluated by using ROIs placed on the phantom core (ROI 1) and the flesh (ROI 2) and averaging their values on the two repetitions (Supplemental Figure 3). The noise was computed as the standard deviation of the subtraction of these two repetitions. Thus, apparent SNR was defined by the following equation:

$$\text{apparent } SNR_{\text{subtraction}} = \frac{\text{mean}(SI_{\text{image1}}, SI_{\text{image2}})}{\text{std}(SI_{\text{image1}} - SI_{\text{image2}})}$$

Accordingly, the apparent SNR, and CNR were measured as follows:

$$\text{apparent } SNR = \text{mean}(\text{apparent } SNR_{\text{ROI1}}, \text{apparent } SNR_{\text{ROI2}})$$

$$CNR = \frac{\text{mean}(SI_{\text{ROI1}}) - \text{mean}(SI_{\text{ROI2}})}{\text{mean}(\text{Noise}_{\text{ROI1}}, \text{Noise}_{\text{ROI2}})}$$



Supplemental Figure 5

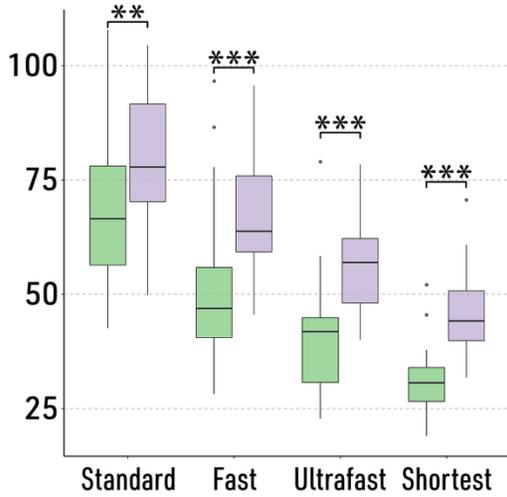
The phantom study used a melon to compute the apparent SNR and contrast-to-noise ratio. Two ROIs were placed close to each other to avoid different receiver weights. The noise was computed from the subtraction image between the same sequences repeated twice.

4.3 Results of apparent SNR and CNR

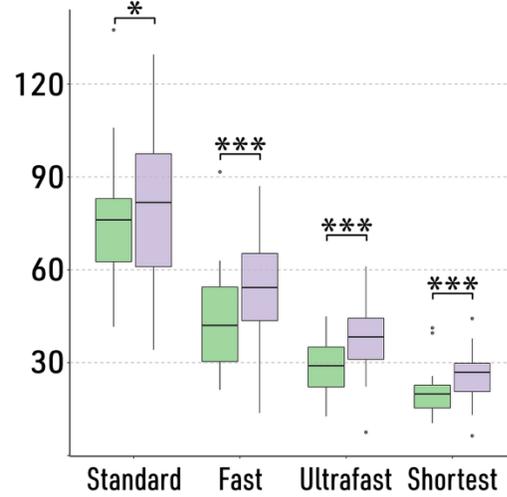
To understand what could drive the effect of dDLR on the detection of contours and lesions, we looked at apparent SNR and lesion CNR in the different conditions. In MS patients, dDLR significantly improved the apparent SNR in all sequences ($p < 0.003$, $p < 0.001$, $p < 0.001$, $p < 0.001$ for standard, fast, ultrafast, and shortest FLAIR, respectively; Supplemental Fig. 4-A). The effect on CNR was also apparent ($p = 0.036$, $p < 0.001$, $p < 0.001$, $p < 0.001$ for standard, fast, ultrafast, and shortest FLAIR, respectively).

The dDLR effect on the phantom was in accordance with previous results (Supplemental Fig. 4-B). Both apparent SNR and CNR were improved in all FLAIR sequences, and the effect was more substantial on rapid sequences. Although a statistical comparison was not conducted for the fruits phantom because of the small number of repetitions, the apparent SNR of ultrafast FLAIR with dDLR reached close level compared to standard FLAIR.

A Patient apparent SNR

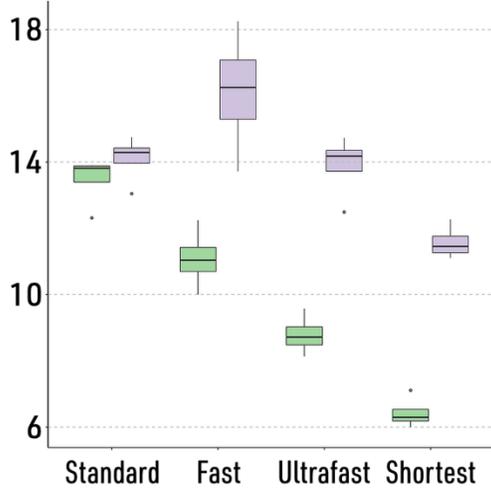


CNR

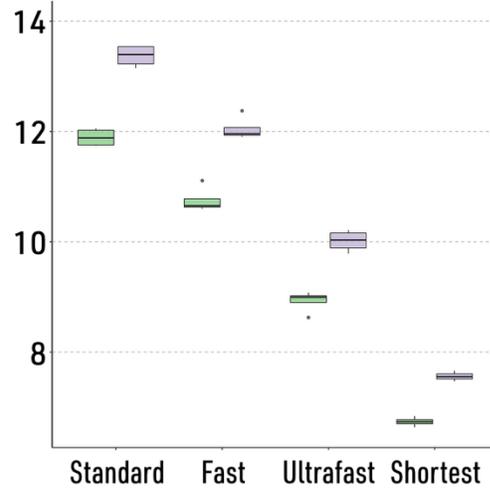


Deep Learning Reconstruction ■ No ■ Yes

B Phantom apparent SNR



CNR



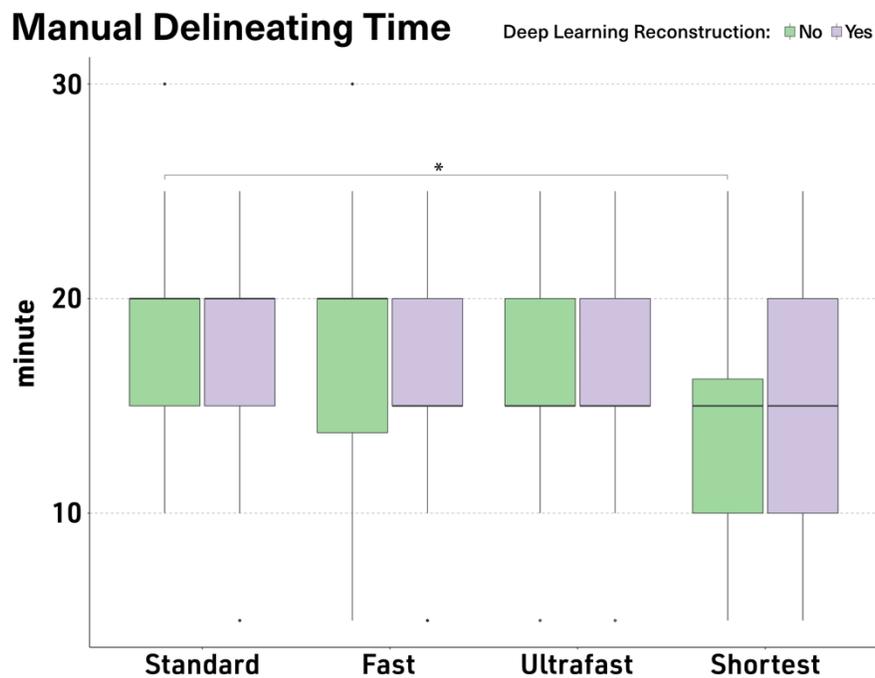
Supplemental Figure 6

(A) Apparent signal-to-noise ratio (SNR), lesion contrast-to-noise ratio (CNR), and lesion contrast resolution (CR) in the patient images and (B) apparent SNR, CNR, CR in the phantom study. While there were decreasing trends of apparent SNR and CNR as the scan time shortened, dDLR significantly improved both.

*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

5. Time for the manual delineation of lesions

Manual delineation was performed randomly and blindly based on the type of FLAIR sequence and the type of reconstruction, but the reader could recognize this. To ensure that this could not indirectly bias the way of contouring, we computed the time spent delineating each sequence. It took an average of 16.4 ± 5.7 minutes for the reader to draw the contours of all the lesions for a single sequence. The time spent contouring slightly decreased along the type of FLAIR sequence, but there was only a significant difference between standard and shortest FLAIR ($p=0.022$) and no significant effect of dDLR (Supplemental Fig. 4). This result ensures that all the images were considered equally. For statistics, repeated two-way ANOVA was used followed by the post hoc test using Tukey Honest Significant Differences.



Supplemental Figure 7

Box plots of time spent in manual delineation. The delineation time was within the same range and only showed a slight decrease for the shortest FLAIR images

* indicates $p < 0.05$.