Differences in CT Perfusion Summary Maps for Patients with Acute Ischemic Stroke Generated by 2 Software Packages


AJNR Am J Neuroradiol 2012, 33 (11) 2074-2080
doi: https://doi.org/10.3174/ajnr.A3110
http://www.ajnr.org/content/33/11/2074
S


troke is the third most common cause of death and the most common cause of disability in the Western world. CTP is emerging as a promising diagnostic tool for the initial evaluation of patients with acute ischemic stroke. In CTP images, areas with perfusion defects can be detected immediately after the onset of clinical symptoms.

CTP analysis results in brain perfusion maps indicating several parameters: CBV, CBF, MTT, and TTP. These parameters are combined in a summary map to quantitatively determine the area of infarct core (sometimes referred to as NVT) and the area of penumbra (sometimes referred to as TAR). Previous studies have shown that the estimation of the size of infarct core and penumbra area is valuable information for predicting the benefit of treatment. However, before such a new analysis is adapted in clinical practice, sufficient evidence of its robustness and accuracy should be provided. The CTP analysis might be influenced by both vendor-specific hardware for CTP image acquisition and software CTP analysis settings and algorithms.

With the increasing availability of quantitative CTP analysis software, it becomes important to understand the potential pitfalls. It has been shown that differences in CTP hardware and software can affect the results. Additional known pitfalls of CTP analysis include incorrect placement of the perfusion volume, incorrect selection and variability of the AIF and VOF, chance of missing small infarcts due to the low resolution of CTP analysis, and changes in perfusion due to extracranial and intracranial stenosis. On the other hand, it has recently been shown that despite the general belief, the order of scanning (CTA before or after CTP) has no significant influence on quantitative CTP parameters.

Although commercial software packages for CTP analysis are widely available, there is currently no standardized method for the analysis. Several algorithms have been developed, applying different perfusion models. Kudo et al demonstrated in 10 patient image datasets that brain perfusion maps resulting from CTP analysis by 5 different commercial software packages (GE Healthcare, Philips Healthcare, Siemens, Toshiba, and Hitachi) may vary considerably. The algorithms of these software packages were categorized into 2 groups on the basis of the applied model and the effect of the delay of the bolus tracer: delay-sensitive and delay-insensitive. A more recent study showed that intervendor differences constituted the primary cause of the variability in CTP analysis results in a population of 11 patients.

In this study, we assessed quantitative differences in CTP summary maps between 2 software packages analyzing identical CTP source images. These 2 software packages represent the 2 mainstream algorithms: delay-sensitive and delay-insensitive. We compared CTP analysis results produced by Extended Brilliance Workspace (Philips Healthcare, Best, the Netherlands) (package A), which represents the delay-sensitive algorithm; and syngo (Siemens, Erlangen, Germany) (package B), which represents the delay-insensitive algorithm.

To our knowledge, this is the first study that reports the variability of CTP summary maps of commercially available
software packages. A summary map uses CBV, CBF, and MTT maps in a single depiction that quantitatively describes both infarct core and penumbra area. Such summary maps are now commonly presented in commercial software packages. These summary maps, rather than the primary perfusion parameters, have the potential to become a major determinant of stroke management. On the basis of a larger patient population than was analyzed in previous studies, we additionally calculated the correlation of infarct core and penumbra area estimated between the 2 algorithms.

Materials and Methods

Patient Population

CTP image data of patients suspected of having acute ischemic stroke were retrospectively collected from February 2010 to March 2011. All datasets of patients with a section thickness of 9.6 mm were included. Exclusion criteria were the following: severe motion artifacts, patients with previous craniotomy, and patients with poor cardiac function. Permission from the medical ethics committee was given for this retrospective analysis of anonymous patient data. Informed consent was waived because no diagnostic tests other than routine clinical imaging were used in this study. Because the results of the evaluation of the images for the purpose of the current study were performed retrospectively, they could not influence clinical decisions.

Imaging Protocols

All scans were performed on a 64-section scanner (Somatom Sensation 64; Siemens). Forty milliliters of iopromide (Ultravist 320; Bayer HealthCare Pharmaceuticals, Pine Brook, New Jersey) was infused at 4 mL/s by using an 18-ga cannula in the right antecubital vein. Acquisition and reconstruction parameters were as follows: 80-kV tube voltage, 150 mAs, collimation of 24 × 1.2 mm, FOV = 300 mm, reconstructed section width of 9.6 mm. At the level of the third ventricle, every 1.5 seconds, images were acquired for the first 50 seconds, followed by a 4-minute lasting image acquisition every 30 seconds. Subsequently, at the level of the roof of the lateral ventricles, only a 50-second lasting acquisition with imaging every 1.5 seconds was performed.

CT Perfusion Analysis

The CTP image data were analyzed by a single trained author (F.F.) blinded from all clinical data by using 2 software packages: Extended Brilliance Workspace 4535 674 25061, Version 3.5, Brain CT Perfusion Package (Philips Healthcare) (package A); and syngo, Version CT 2007A, Neuro Perfusion CT package (Siemens) (package B).

The postprocessing steps conducted for both software packages in the process of registration, segmentation, and perfusion parameter definition were inspected by an experienced radiologist (L.B.) to ensure identical input parameters for CTP analysis. The input parameters included the AIF, VOF, hematocrit, CBV threshold, CBF threshold, and relative MTT threshold. The arterial input function is required to perform a deconvolution with the time-intensity curves of the brain tissue. The venous output function is required to correct the arterial input for volume-averaging effects. The hematocrit is the ratio of red blood cell volume to the total volume of blood. This factor is used to convert contrast enhancement information (in Hounsfield units) to mass density.
units) to CBV in milliliters/100 g of tissue. It should not be adjusted without actually measuring the patient’s hematocrit and accordingly was set at the default value of 0.45. Software package A selects the whole ischemic area on the basis of a relative MTT threshold, defined as the area in which the MTT is increased 1.5 times compared with the contralateral side. Software package B uses the CBF threshold to identify the areas of perfusion abnormality. Both software packages use the CBV threshold to identify what part of the area of perfusion abnormality is salvageable (penumbra) or not (infarct core).

Both software packages include automatic registration of the images, which was not manually modified. By default, cerebral segmentation was also automatically performed in both software packages with vendor default threshold values. If a manually generated mask was required, a similar cerebral area was generated for both software packages.

Because the AIF and VOF are generated semiautomatically in software package B, this analysis was performed first. Most commonly, the AIF was determined in an anterior cerebral artery. The VOF was generally determined in the superior sagittal sinus.3,14 For software package A, the same region of interest was chosen for the AIF and VOF generation as in software package B. The maximum intensity in Hounsfield units and time to peak of the AIF and VOF were inspected to ensure similarity. The midlines were manually set to be the same for both packages. Subsequently, the summary maps were generated by using the same settings, a hematocrit of 0.45 and vessel removal at the threshold value of 9 mL/100 g in the CBV map.

The labeling of the voxels as infarct core and penumbra was performed by using the factory settings of the thresholds. In software package A, the infarct core was defined as pixels with a measured relative MTT > 1.5 and measured CBV < 2.0 mL/100 g; the area of penumbra was defined with a measured relative MTT > 1.5 and a measured CBV > 2.0 mL/100 g. In software package B, the infarct core is called NVT and was defined with a measured CBF < 20 mL/100 g/min and a measured CBV < 2.0 mL/100 g. The equivalent of the penumbra, TAR was defined with a measured CBF < 20 mL/100 g/min and a measured CBV > 2.0 mL/100 g.

From this point, NVT and TAR are referred to as infarct core and penumbra, respectively. In the summary maps, the infarct core was presented in red. The penumbra was displayed in green and in yellow for software packages A and B, respectively (Fig 1). To study whether potential differences are caused by using the MTT threshold versus the CBF threshold, we also analyzed the summed area of infarct core and penumbra. We defined this summed area as area of perfusion abnormality. For software package A, this is equal to the area of relative MTT > 1.5; for package B, this is the area with a CBF < 20 mL/100 g/min.

**Statistical Analysis**

Means and SDs of absolute and relative differences of the area of infarct core, penumbra, and perfusion abnormality as determined by the 2 software packages were calculated. The relative difference of the measured area was defined as the ratio of the difference in area to the average area and was represented as a percentage. The relation between the measurements based on both software packages was assessed with scatterplots and with the calculation of linear regression lines. Correlation between the values was evaluated by calculating the Pearson correlation. Agreement between the 2 software packages was tested by calculating the systematic error (bias) and the 95% limits of agreement, defined as the bias ± 1.96 of the individual differences, as part of a Bland-Altman analysis. The dependency between the 2 methods was tested by linear regression of differences as shown in the Bland-Altman plots. If the 2 methods are equally variable, the slope of this linear regression line would equal zero. P values smaller than .05 were considered statistically significant.

**Results**

From the 30 consecutive patients, 26 were included and 41 generated image datasets were used. As part of the protocol of the multicenter Dutch Acute Stroke Trial (http://www.dutchstroketrial.nl), some patients have had follow-up CTP. One patient was excluded because of a craniotomy, and 3 patients were excluded because of severe motion artifacts. Of the 26 patients, the average age was 58 years, ranging from 26 to 91 years; 16 were men.

Figure 2 shows a typical example of a summary map from identical patient data generated by the 2 packages. This figure shows a similar area of perfusion abnormalities, but a different size of infarct core and penumbra. The Table shows the average of the absolute differences in area measurements by the 2 software packages. There was a significant absolute difference in the area of infarct core and penumbra between the 2 packages. The relative difference in infarct core and perfusion ab-
normality was also significant. The scatterplots of the area of infarct core, penumbra, and perfusion abnormality of the summary maps are shown in Fig 3. The correlation coefficient of the area of infarct core between the software packages was $r = 0.62$ ($P < .001$); the regression resulted in a slope of 0.78 ($P < .001$). For the area of the penumbra, the correlation coefficient was $r = 0.28$ ($P = .07$); the slope of the regression line was 1.26 ($P = .07$). The area of perfusion abnormality had a better correlation coefficient of $r = 0.70$ ($P < .001$). The linear regression generated from this relation had slope of 1.02 ($P < .01$).

Figure 4 illustrates the Bland-Altman analysis for the area measurements. A statistically significant regression line was observed only for the penumbra area and area of perfusion abnormality. The limits of agreement for each measurement of the infarct core, penumbra, and perfusion abnormality area are shown in Fig 4 and are given in the Table. The Bland-Altman limits of agreement were $-74.8–27.6$ cm$^2$ for the infarct core area, $-34.7–66.3$ cm$^2$ for penumbra area, and $-67.0–51.4$ cm$^2$ for perfusion abnormality area.

### Discussion

We found large differences in estimated infarct core and penumbra areas resulting from the 2 perfusion CT software packages. The Bland-Altman analysis showed severe lack of agreement reflected by the large intervals of agreement for each measurement, with discrepancies of $>100$ cm$^2$.

In general, analysis with software package A resulted in larger penumbra areas, and analysis with software package B, in larger infarct core areas. There were many cases for which software package A estimated a small area of infarct core (smaller than 10 cm$^2$), and software package B estimated a large infarct core (up to 80 cm$^2$, Fig 3A). On average, the infarct core area was 30% smaller for software package A, and the penumbra area was approximately 30% larger. The average area of perfusion abnormalities was not significantly different, but there was a large spread of the differences. This was also illustrated in the typical example of a difference in infarct core and penumbra area, with a similar area of perfusion abnormality (Fig 2).

The correlation between both software packages for the penumbra area was especially weak. A higher correlation coefficient was observed in the area of perfusion abnormality. The linear correlation for both the penumbra area and perfusion abnormality area in the Bland-Altman analysis of Fig 4 revealed a dependency between the difference of the 2 methods and their average. Such a dependency is probably due to a systematic difference between these methods.

The 2 software packages use a different algorithm for the CTP analysis. Software package A uses a delay-sensitive algo-
Different mathematic methods for CTP analysis were described by Wintermark et al.\textsuperscript{9} One approach uses a deconvolution model; the other uses a nondeconvolution maximum slope model. In software package A, CBV, CBF, and MTT maps were calculated by using a deconvolution algorithm. Deconvolution is a mathematic process that compensates the effects of the AIF on the time-attenuation curve to calculate the perfusion parameters. In software package B, the maximum slope technique was used, which takes the slope of bolus arrival time and the maximum value of the time-attenuation curve into account and, therefore, is considered delay-insensitive. Therefore, CTP parameters based on both methods may have the same name but are actually defined and calculated quite differently. Kudo et al.\textsuperscript{11} have already shown that these differences in the CTP parameter maps are considerable.

This difference in algorithms used in both software packages is crucial in generating the CTP maps of CBV, CBF, and MTT.\textsuperscript{11} Furthermore, there is also a difference in the way both software packages define the infarct core and penumbra on the basis of generated CTP maps. In software package A, the infarct core is defined as the area with a relative MTT value 50% higher than that in the other hemisphere (ie, relative MTT > 1.5) and a CBV value lower than 2.0 mL/100 g. If the CBV value is larger than this threshold, the area is defined as penumbra. Software package B defines infarct core as the area with a CBF value below 20 mL/100 g/min and a CBV value as below 2.0 mL/100 g. If the CBV value is larger than this threshold the area is defined as penumbra. We suspected that this difference in definition also contributes to the large differences in area estimated in this study.

For decades, reduced CBF has been associated with ischemia. Because the flow cannot be directly measured from CTP data, it has to be estimated by using dynamic parameters such as MTT, which also represent flow defects because it equals volume/flow by definition. However, both parameters have different dimensions and scaling. The concern is how to estimate MTT as well as CBF. Both vendors use algorithms for this estimation that lead to quite different results. This does not disqualify contrast dynamics techniques but should inspire us to identify optimal ways to estimate local perfusion.

Our findings agree with those of Kudo et al.\textsuperscript{11,12} who studied 5 commercially available software packages and classified these into 2 groups based on differences in tracer-delay sensitivity. These authors found that CTP maps correlated well within the classified groups but not across them. The packages used in the current study are from these 2 separate groups. They showed that a delay-sensitive algorithm has the tendency to produce substantial differences in CBF and MTT, with a decrease in CBF and an increase in MTT for positive delays and vice versa for negative delays. This affected the estimation of infarct core and penumbra area.\textsuperscript{12,15} Another study showed that delay-sensitive algorithms may overestimate CBV values in patients with concomitant intra- or extracranial severe hemodynamic delays.\textsuperscript{16} Adjusting the threshold value for CBV in software package A might result in a better correspondence with software package B. A recent study reported that it is CBF—not CBV—that has highest accuracy compared with DWI as the criterion standard in defining infarct core.\textsuperscript{17} In that study, it was shown that adjusting the threshold value for CBF resulted in a better correlation with DWI. Yet, again, the adjustment of CBF threshold differed for each software package.\textsuperscript{17}

CTP analysis results are sensitive to several parameters. Preceding studies revealed that the CTP analysis results may also vary due to scan parameters\textsuperscript{18} and postprocessing steps such as the defining of input and output function.\textsuperscript{19} Other studies also assessed the reproducibility of CTP due to inter-

---

**Fig 4.** Bland-Altman plot including the limits of agreement between the 2 software packages in the infarct core (A), penumbra (B), and area of perfusion abnormality (C). For Bland-Altman plots, only the significant linear regression lines are shown. For the penumbra area, the slope is 1.62 ($P < .001$); for the area of perfusion abnormality, the slope is 0.43 ($P < .001$).
and intrauser variability as the result of a different selection of parameters.20,21 Even though there was a high degree of correlation between and within users in producing the CBV, CBF, and MTT maps within a single analysis software package (GE Healthcare), the level of agreement was considered not sufficient to allow quantitative data derived from these maps for clinical decision-making.21 Recently, it was shown that intra-vendor differences are the primary cause of the variability in CTP analysis.14

In this study, we have assessed the reproducibility of the CTP analysis summary maps between both methods, rather than the accuracy of either method. We, therefore, abstain from any judgment on which package is more accurate. Evaluating the accuracy of the CTP (summary) maps is a difficult task. CTP results can be compared with DWI, which is the current widely accepted de facto clinical reference standard for the determination of the infarct core.22,23 DWI is, however, not widely available in the acute setting, and during the time between CTP imaging and DWI, the infarct core could increase. Furthermore, the assessment of the penumbra with DWI is more difficult. The study of Kamalian et al17 showed that the optimization of perfusion parameter thresholds to obtain the best agreement with DWI was also dependent on the analysis software package. The accuracy of the CTP summary map was also addressed by Wintenmark et al,24 who compared CTP analysis with follow-up CT or MR imaging and showed that CTP-based analyses are more accurate in detecting hemispheric stroke than using admission nonenhanced CT. However, analysis of the accuracy of the actual area measurements of infarct core and penumbra was not performed in this study. Also, software packages are introduced in the market without published clinical validations of the measurements. Often the algorithms of these packages are not published either. As a result, physicians may view these software packages as “black boxes.”

There were several limitations to this study. First, the summary maps that have been studied here are derived from the vendor-specific CBV, CBF, MTT, and TTP estimations. It is expected that the differences in these maps will result in differences in the infarct core and penumbra area in the summary maps. Differences in CBV, CBF, and MTT have already been addressed in previous studies11,12 and were not the subject of this study. Instead, we studied how these differences in combination with a different definition of infarct core and penumbra area result in biases of infarct core and penumbra area. Because detail of the algorithms of the software packages was restricted by the vendors, we were not able to present a detailed explanation of the origin of differences in the summary map. A second limitation is that software packages for CTP analysis may be prone to updates of the algorithms and the used packages may already have been changed, resulting in different outcomes as presented here. Finally, not all parameters were completely identical for the 2 software packages because it was not possible to adjust these. We tried to optimize the similarity, but some processes could not be truly identical. For example, the software packages used an internal registration of the CTP images, which was difficult to adjust.

We only studied the variability of CTP analysis resulting from using different software. Both vendors also provide CT scanners, which may vary as well in the generation of CTP image data, either due to scanner hardware or image-reconstruction parameters. Because these scanner systems are regularly calibrated, we expect that this variation is smaller than the variation due to using different software-analysis packages.25 However, this was not investigated in this study. Such a comparison would require additional CT scanning of patients with stroke and is, therefore, unethical because of an increase of radiation and contrast material and a delayed treatment.

Conclusions
We observed large differences in the summary maps generated by 2 software packages, representing the 2 main types of CTP analysis. In our study of 41 cases, the differences in infarct core and penumbra area were statistically significant and the degree of agreement was not acceptable. Because of this variability, CTP summary maps should be interpreted with care. This study emphasizes the need for standardization of CTP analysis algorithms, and further research and protocol development are advocated before CTP can become a robust determinant of stroke management in clinical practice.

Acknowledgements
This work has been supported by LP3M University of Sumatera Utara and USU Academic Hospital through Directorate General of Higher Education (DIKTI), Ministry of National Education Indonesia.


References


