

# Diffusion-Weighted Imaging of the Head and Neck in Healthy Subjects: Reproducibility of ADC Values in Different MRI Systems and Repeat Sessions

A.S. Kolff-Gart, P.J.W. Pouwels, D.P. Noij, R. Ljumanovic, V. Vandecaveye, F. de Keyzer, R. de Bree, P. de Graaf, D.L. Knol, and J.A. Castelijns



## ABSTRACT

**BACKGROUND AND PURPOSE:** DWI is typically performed with EPI sequences in single-center studies. The purpose of this study was to determine the reproducibility of ADC values in the head and neck region in healthy subjects. In addition, the reproducibility of ADC values in different tissues was assessed to identify the most suitable reference tissue.

**MATERIALS AND METHODS:** We prospectively studied 7 healthy subjects, with EPI and TSE sequences, on 5 MR imaging systems at 3 time points in 2 institutions. ADC maps of EPI (with 2 b-values and 6 b-values) and TSE sequences were compared. Mean ADC values for different tissues (submandibular gland, sternocleidomastoid muscle, spinal cord, subdiaphragmatic lymph node, and tonsil) were used to evaluate intra- and intersubject, intersystem, and intersequence variability by using a linear mixed model.

**RESULTS:** On 97% of images, a region of interest could be placed on the spinal cord, compared with 87% in the tonsil. ADC values derived from EPI-DWI with 2 b-values and calculated EPI-DWI with 2 b-values extracted from EPI-DWI with 6 b-values did not differ significantly. The standard error of ADC measurement was the smallest for the tonsil and spinal cord (standard error of measurement =  $151.2 \times 10^{-6} \text{ mm}^2/\text{s}^2$  and  $190.1 \times 10^{-6} \text{ mm}^2/\text{s}^2$ , respectively). The intersystem difference for mean ADC values and the influence of the MR imaging system on ADC values among the subjects were statistically significant ( $P < .001$ ). The mean difference among examinations was negligible (ie,  $<10 \times 10^{-6} \text{ mm}^2/\text{s}^2$ ).

**CONCLUSIONS:** In this study, the spinal cord was the most appropriate reference tissue and EPI-DWI with 6 b-values was the most reproducible sequence. ADC values were more precise if subjects were measured on the same MR imaging system and with the same sequence. ADC values differed significantly between MR imaging systems and sequences.

**ABBREVIATIONS:** 2b = 2 b-values; 6b = 6 b-values; IQR = interquartile range; LoA = limits of agreement; SEM = standard error of measurement

Almost 3% of all malignancies are head and neck cancer, 95% of which are squamous cell carcinomas.<sup>1</sup> MR imaging is one of the modalities used in the work-up of patients with head and neck cancer.<sup>2</sup> DWI is an MR imaging technique by which diffusion properties of water can be quantified as an ADC value.<sup>3</sup> Changes in ADC are inversely correlated with changes in cellularity.<sup>4</sup> In tissues with high cellularity, diffusion of extracellular water in particular is limited by cell membranes, which

give low ADC values. In tissues with low cellularity, when diffusion is facilitated (eg, in edematous or necrotic tissue), ADC values are high.

Indications for DWI in head and neck cancer include tissue characterization of primary tumors and nodal metastases, prediction and monitoring of treatment response after chemotherapy or radiation therapy, and differentiation of radiation changes and residual or recurrent disease.<sup>5</sup>

Neither the optimal DWI sequence for assessment of the head and neck region nor its reproducibility has been clearly established, to our knowledge. DWI can be performed with either EPI or TSE sequences, of which the EPI sequence is most commonly used in the head and neck area.<sup>6,7</sup> On EPI-DWI, more malignant lesions can be detected and lesion delineation is facilitated. However, the interobserver agreement of ADC values is reported to be higher on TSE-DWI, probably due to the frequent occurrence of artifacts and geometric distortions in EPI-DWI.<sup>8</sup>


Currently the use of DWI in head and neck imaging is mostly confined to research protocols and advanced academic centers.

Received May 6, 2014; accepted after revision July 28.

From the Departments of Radiology and Nuclear Medicine (A.S.K.-G., D.P.N., R.L., P.d.G., J.A.C.), Physics and Medical Technology (P.J.W.P.), Otolaryngology-Head and Neck Surgery (R.d.B.), and Clinical Epidemiology and Biostatistics (D.L.K.), VU University Medical Center, Amsterdam, the Netherlands; and Department of Radiology (V.V., F.d.K.), University Hospitals Leuven, Leuven, Belgium.

Please address correspondence to Jonas A. Castelijns, MD, PhD, Department of Radiology and Nuclear Medicine, VU University Medical Center, PO Box 7057, 1007 MB Amsterdam, the Netherlands; e-mail: j.castelijns@vumc.nl

 Indicates article with supplemental on-line appendix.

 Indicates article with supplemental on-line photo.

<http://dx.doi.org/10.3174/ajnr.A4114>

Before DWI can be used in multicenter studies, its reproducibility across different centers and MR imaging systems should be validated.<sup>9</sup> ADC values may be affected by the selected technique and MR imaging system (eg, due to differences in gradient systems, coils, pulse-sequence designs, imaging parameters, and artifacts related to susceptibility effects or eddy currents).<sup>10</sup> Information on variance is needed.<sup>11</sup> Furthermore, the use of reference tissues might help ascertain the variability among different MR imaging systems and could potentially help correct for differences in ADC values among MR imaging systems.

The purpose of this prospective study was to determine the reproducibility of ADC values in the head and neck region obtained from DWI on the basis of both EPI and TSE sequences in repeated measurement on different MR imaging systems in healthy subjects. In addition, we assessed which tissue shows the highest reproducibility in ADC values, so that it could function as a reference tissue in future studies.

## MATERIALS AND METHODS

### Subjects

The study population consisted of 7 healthy subjects, 5 men and 2 women (age range, 27–54 years; median age, 30 years). The subjects were examined in 2 institutions: VU University Medical Center and University Hospitals Leuven. All examinations were performed in 2011, after obtaining approval from the relevant institutional review boards and written informed consent from all subjects. We used the

following MR imaging systems: I) Avanto (Siemens, Erlangen, Germany), II) Sonata (Siemens), III) Signa HDxt (GE Healthcare, Milwaukee, Wisconsin), and IV) Aera (Siemens), all at 1.5T, and V) Achieva (Philips Healthcare, Best, the Netherlands) at 3T. All examinations were performed with a dedicated head and neck radiofrequency coil in combination with a spine-array coil.

All subjects were examined on all MR imaging systems at 3 time points per MR imaging system, yielding a total of 15 sessions per subject. Two examinations were performed on the same day (between examinations, the subject was removed from the MR imaging system), and 1 examination, at least 1 month later.

### Imaging Protocol

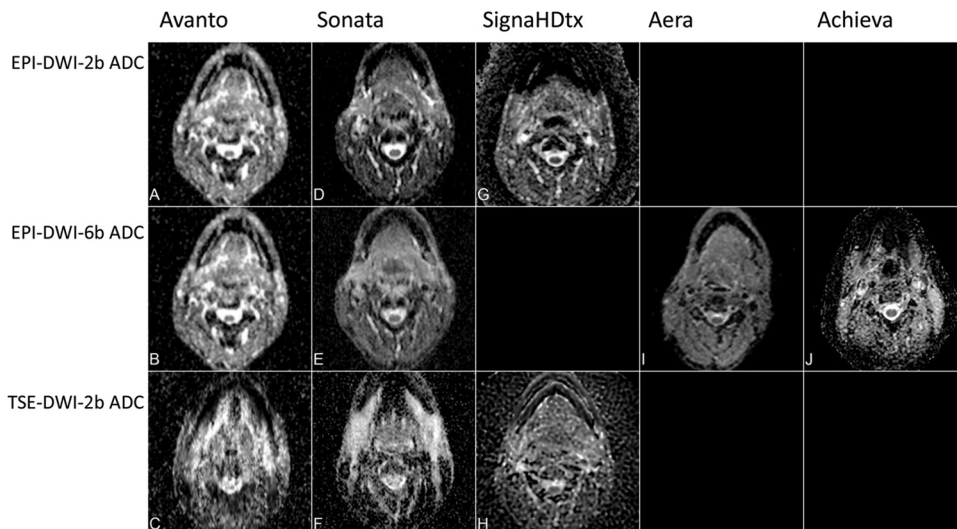
Each session included an anatomic T2-weighted sequence through the neck and up to 3 DWI sequences, with acquisition parameters as similar as possible among the MR imaging systems. Due to technical limitations, no EPI-DWI with 6 b-values (6b) was performed on 1 MR imaging system (Signa HDxt), and on 2 MR imaging systems (Aera and Achieva), no separate EPI-DWI with 2 b-values (2b) was performed. The sequences used per MR imaging system are shown in Table 1 and Fig 1.

All imaging was acquired with 21 transverse sections centered on the epiglottis (section thickness, 4 mm; intersection gap, 0.4 mm). The imaging protocol consisted of both conventional T2-weighted (TR/TE, at least 3700/90–110 ms; in-plane pixel size,  $0.95 \times 0.95$  mm) and EPI-DWI (TR/TE, at least 4300/59–98 ms; in-plane pixel size,  $1.5\text{--}1.9 \times 1.5\text{--}1.9$  mm; interpolated in-plane pixel size,  $0.75\text{--}0.95$  mm) or TSE-DWI (TR/TE, 900–3000/84–113 ms; in-plane pixel size,  $1.3 \times 1.3$  mm). B-values for the 3 DWI series were as follows: 1) EPI-DWI obtained with 6 b-values of 0, 50, 100, 500, 750 and 1000  $\text{s}/\text{mm}^2$ , 2) EPI-DWI obtained with 2 b-values of 0 and 1000  $\text{s}/\text{mm}^2$ , and 3) TSE-DWI obtained with 2 b-values also of 0 and 1000  $\text{s}/\text{mm}^2$ .

**Table 1: Specification of DWI sequences obtained at each MRI system**

|                 | I         | II        | III        | IV      | V       |
|-----------------|-----------|-----------|------------|---------|---------|
| Manufacturer    | Siemens   | Siemens   | GE         | Siemens | Philips |
| Model           | Avanto    | Sonata    | Signa HDxt | Aera    | Achieva |
| Center          | Amsterdam | Amsterdam | Amsterdam  | Leuven  | Leuven  |
| Field strength  | 1.5T      | 1.5T      | 1.5T       | 1.5T    | 3T      |
| Conventional T2 | +         | +         | +          | +       | +       |
| EPI-DWI-2b      | +         | +         | +          | o       | o       |
| EPI-DWI-6b      | +         | +         | –          | +       | +       |
| TSE-DWI-2b      | +         | +         | +          | –       | –       |

**Note:**—Amsterdam indicates VU University Medical Center; Leuven, University Hospitals Leuven; +, sequence performed; –, sequence not performed, not available; o, data extracted from 6b.

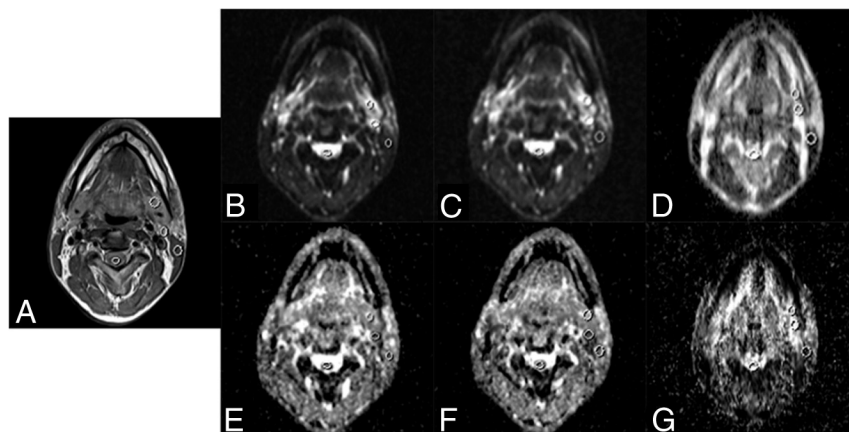


**FIG 1.** ADC maps of all DWI sequences on all MR imaging systems. On Signa HDxt (GE Healthcare), EPI-DWI-6b ADC was not performed. On Aera (Siemens) and Achieva (Philips Healthcare), EPI-DWI-2b ADC was extracted from EPI-DWI-6b and TSE-DWI-2b was not performed.

**Table 2: Number of placed ROIs per tissue and per sequence<sup>a</sup>**

|                      | Tonsil   | Spinal Cord | SCM      | SMG      | SDG LN   | Total     |
|----------------------|----------|-------------|----------|----------|----------|-----------|
| EPI-DWI-2b (No.) (%) | 58 (90)  | 74 (96)     | 75 (97)  | 76 (99)  | 74 (99)  | 357 (96)  |
| EPI-DWI-6b (No.) (%) | 57 (89)  | 76 (99)     | 76 (99)  | 76 (99)  | 74 (99)  | 359 (97)  |
| TSE-DWI-2b (No.) (%) | 50 (83)  | 76 (97)     | 77 (97)  | 76 (96)  | 72 (97)  | 351 (95)  |
| Total                | 165 (87) | 226 (97)    | 228 (98) | 228 (98) | 220 (98) | 1067 (96) |

**Note:**—SCM indicates sternocleidomastoid muscle; SMG, submandibular gland; SDG LN, subdigastic lymph node.  
<sup>a</sup> The percentage of the maximum number of possible ROIs displayed is in parentheses. Elimination is due to poor image quality or artifacts.



**FIG 2.** Examples of regions of interest drawn on T2 (A), EPI-DWI-2b B0 (B), EPI-DWI-6b B0 (C), TSE-DWI-2b B0 (D), EPI-DWI-2b ADC (E), EPI-DWI-6b ADC (F), and TSE-DWI-2b ADC (G). The tonsils are not visible at this level. Images were acquired with Avanto (Siemens).

### Data Analysis

All ADC maps were calculated on-line or off-line by using the MR imaging system software of the respective vendor. EPI-DWI-6b was analyzed assuming a monoexponential ADC. ADC values for EPI-DWI-2b on the 2 MR imaging systems without EPI-DWI-2b were derived from EPI-DWI-6b by selecting only the images acquired by using  $b=0$  s/mm<sup>2</sup> and  $b=1000$  s/mm<sup>2</sup>.<sup>12</sup> These “generated” EPI-DWI-2b data were compared with the other EPI-DWI-2b data. Data were transferred to a DICOM viewer.

For each examination, 1 elliptic region of interest per tissue was manually drawn on the section that contained the bulk of the tissue of interest by 1 observer (R.L.) with 7 years of experience in head and neck imaging. ADC values were determined for each of the following 5 tissues in the head and neck: 1) submandibular gland, 2) sternocleidomastoid muscle, 3) spinal cord, 4) subdigastic lymph node, and 5) tonsil. For the selection of a subdigastic lymph node, either the left or right one was selected consistently within each subject. The size (range, 20–50 mm<sup>2</sup>) and the position of the region of interest were identified on T2-weighted images. ROIs were drawn on corresponding B0 images by visual comparison with the anatomic T2WI. ROIs drawn on  $b=0$  s/mm<sup>2</sup> images were copied to the corresponding ADC maps.

### Statistical Analysis

First, it was determined whether ADC values of the EPI-DWI-2b sequences can be replaced with ADC maps obtained by selecting only the  $b=0$  and  $b=1000$  s/mm<sup>2</sup> images from the EPI-DWI-6b (for MR imaging systems IV and V), because they are theoretically equivalent. We used a linear mixed model, with fixed effects for subjects,

MR imaging systems, sequences, and an MR imaging system  $\times$  sequences interaction.<sup>13,14</sup> Random effects were all possible interactions with the subjects (On-line Appendix). This possibility was tested by using data from MR imaging systems I and II, these being the only MR imaging systems on which both sequences had been performed.

For the main variance analysis, 5 MR imaging systems and 3 sequences were compared by using the same statistical modeling approach and reasoning as those used for the linear mixed model and by incorporating tissues as fixed effects (On-line Appendix). All 3 examinations of each subject were assumed pure replications and were nested within subject  $\times$  MR imaging system combinations. Models with sequence-specific error variances were compared by using the Akaike Information Criterion.<sup>15</sup> The standard error of measurement (SEM) for ADC values per tissue was expressed as the square root of the sum of residual variance ( $\sigma^2_E$ ) and the variance expressing the interaction between replication and subjects at different

MR imaging systems ( $\sigma^2_{R:IM}$ ), sequences ( $\sigma^2_{SR:IM}$ ), and tissues ( $\sigma^2_{TR:IM}$ ) (On-line Appendix):

$$SEM = \sqrt{\sigma^2_{R:IM} + \sigma^2_{SR:IM} + \sigma^2_{TR:IM} + \sigma^2_E}$$

Differences in mean ADC values for all systems and the between-subjects effects were tested by using a Levene Test of Equality of Error Variances; an  $\alpha$ -level of .05 was used for statistical significance.<sup>16</sup> All missing data or images with poor quality of DWI were specifically labeled for statistical analysis. Boxplots were created by using SPSS (Version 20.0; IBM, Armonk, New York). All other analyses were performed with Proc NL MIXED of SAS (Version 9.2; SAS Institute, Cary, North Carolina).

## RESULTS

### DWI

All subjects underwent multiple DWI sessions with multiple sequences on all MR imaging systems. For MR imaging system III, EPI-DWI-6b was unavailable; for MR imaging systems IV and V, ADC maps for EPI-DWI-2b were constructed by using only the  $b=0$  and  $b=1000$  s/mm<sup>2</sup> images from the EPI-DWI-6b, yielding a total of 12 DWI sequences per subject (Table 1). Two subjects underwent 2 instead of 3 replications. One subject had prior bilateral tonsillectomies. Therefore, the maximum number of possible ROIs was 1104. For a detailed overview of the number of possible ROIs, see On-line Figure. Further elimination was due to technically failed images and image-specific poor quality, and in 37 cases, it was impossible to place a region of interest: In 95% of tissues, region-of-interest placement was possible on TSE-DWI-2b; in 96%, on EPI-DWI-2b; and in 97%, on EPI-DWI-6b (Table 2). Examples of ADC maps on different MR

imaging systems and sequences are shown in Fig 1. An example of drawn ROIs is shown in Fig 2.

When combining the results of the 3 DWI sequences, region-of-interest placement was possible in 96% of tissues (Table 2). However, in only 87% (range, 83%–90%) of images could a region of interest could be placed on the tonsil. In the other regions, ROIs could be placed in 97%–98% of cases. These data indicate that the tonsil is probably not a good reference tissue for future evaluations.

A variance component analysis was performed for MR imaging systems I and II to test potential differences between ADC values derived from the EPI-DWI-2b sequence and the calculated EPI-DWI-2b extracted from EPI-DWI-6b (Table 3). The lowest bias was found in the subdigastric lymph node ( $0.7 \times 10^{-6} \text{ mm}^2/\text{s}$ ), and the highest bias was found in the tonsil ( $-23.2 \times 10^{-6} \text{ mm}^2/\text{s}$ ). Furthermore, this analysis showed a small range of limits of agreement (LoA) (range,  $-307.0 \times 10^{-6} \text{ mm}^2/\text{s}$  to  $302.4 \times 10^{-6} \text{ mm}^2/\text{s}$ ) for all tissues combined. This finding implies that both ADC values are not significantly different. Therefore, we used calculated EPI-DWI-2b ADC values extracted from EPI-DWI-6b on systems if EPI-DWI-2b was not available for further analysis.

The intersystem difference between the MR imaging systems, with mean ADC values as a dependent variable, was statistically significant ( $P < .001$ ). The influence of the sequence, the MR imaging system, and the interaction between these 2 parameters was significant ( $P = .011$ ). The influence of the MR imaging system on the ADC values among the subjects ( $P < .001$ ) was also significant.

### Main Variance Analysis

For the main analysis, the actual median ADC values and the results of the main variance component analysis per tissue are shown in Table 4 and are visualized in a boxplot (Fig 3). The 3 DWI sequences used showed some difference: The EPI-DWI-6b sequence demonstrated the smallest interquartile range (IQR)

**Table 3: Comparison of ADC values derived from calculated EPI-DWI-2b extracted from EPI-DWI-6b and EPI-DWI-2b for MRI systems I and II**

|             | Bias ( $10^{-6} \text{ mm}^2/\text{s}$ ) | LoA ( $10^{-6} \text{ mm}^2/\text{s}$ ) |
|-------------|--|---|
| Tonsil      | -23.2 (-2.9)                             | -307.0–260.7                            |
| Spinal cord | -12.7 (-1.2)                             | -296.6–271.1                            |
| SCM         | 10.8 (1.1)                               | -273.1–294.6                            |
| SMG         | 18.6 (1.3)                               | -265.3–302.4                            |
| SDG LN      | 0.7 (0.1)                                | -283.2–284.5                            |

**Note:**—SCM indicates sternocleidomastoid muscle; SMG, submandibular gland; SDG LN, subdigastric lymph node.

<sup>a</sup> In parentheses, the bias is displayed as a percentage of the mean ADC from EPI-DWI-2b for MRI systems I and II.

values ( $830\text{--}1217 \times 10^{-6} \text{ mm}^2/\text{s}$ ) and lowest SEM (190.3) in ADC for all tissues. The TSE-DWI-2b sequence demonstrated the broadest IQR ( $812\text{--}1414 \times 10^{-6} \text{ mm}^2/\text{s}$ ) and largest SEM ( $284.5 \times 10^{-6} \text{ mm}^2/\text{s}$ ) for all tissues, while EPI-DWI-2b and EPI-DWI-6b showed a narrower IQR ( $815\text{--}1226 \times 10^{-6} \text{ mm}^2/\text{s}$  and  $830\text{--}1217 \times 10^{-6} \text{ mm}^2/\text{s}$ , respectively) and a smaller SEM ( $216 \times 10^{-6} \text{ mm}^2/\text{s}$  and  $190.3 \times 10^{-6} \text{ mm}^2/\text{s}$ , respectively). Therefore measurements on EPI-DWI-2b and EPI-DWI-6b are more precise. Note that with TSE-DWI-2b, the lowest number of ADCs was available for analysis (95%, Table 2).

The spinal cord and tonsil showed the smallest IQR ( $873\text{--}1100 \times 10^{-6} \text{ mm}^2/\text{s}$  and  $694\text{--}980 \times 10^{-6} \text{ mm}^2/\text{s}$ , respectively) and lowest SEM ( $190.1 \times 10^{-6} \text{ mm}^2/\text{s}$  and  $151.2 \times 10^{-6} \text{ mm}^2/\text{s}$ , respectively) (Table 4 and Fig 4). These tissues have the lowest SEM, indicating that ADC measurements in these tissues are the most precise and the best reproducible ones. However, even though the SEM is low for the spinal cord ( $\text{SEM} = 190.1 \times 10^{-6} \text{ mm}^2/\text{s}$ ), with a median ADC of  $976 \times 10^{-6} \text{ mm}^2/\text{s}$ , the range of normal values is still broad ( $\text{IQR} = 873\text{--}1100 \times 10^{-6} \text{ mm}^2/\text{s}$ ).

Variance caused by time is limited (Fig 5). The mean difference in ADC values of the second examination compared with the first, which were on the same day, was  $6 \times 10^{-6} \text{ mm}^2/\text{s}$  ( $\text{SD} = 310 \times 10^{-6} \text{ mm}^2/\text{s}$ ). For the third examination, 1 month after the first, the mean difference in ADC values was  $-5 \times 10^{-6} \text{ mm}^2/\text{s}$  ( $\text{SD} = 310 \times 10^{-6} \text{ mm}^2/\text{s}$ ) compared with the first measurement.

### DISCUSSION

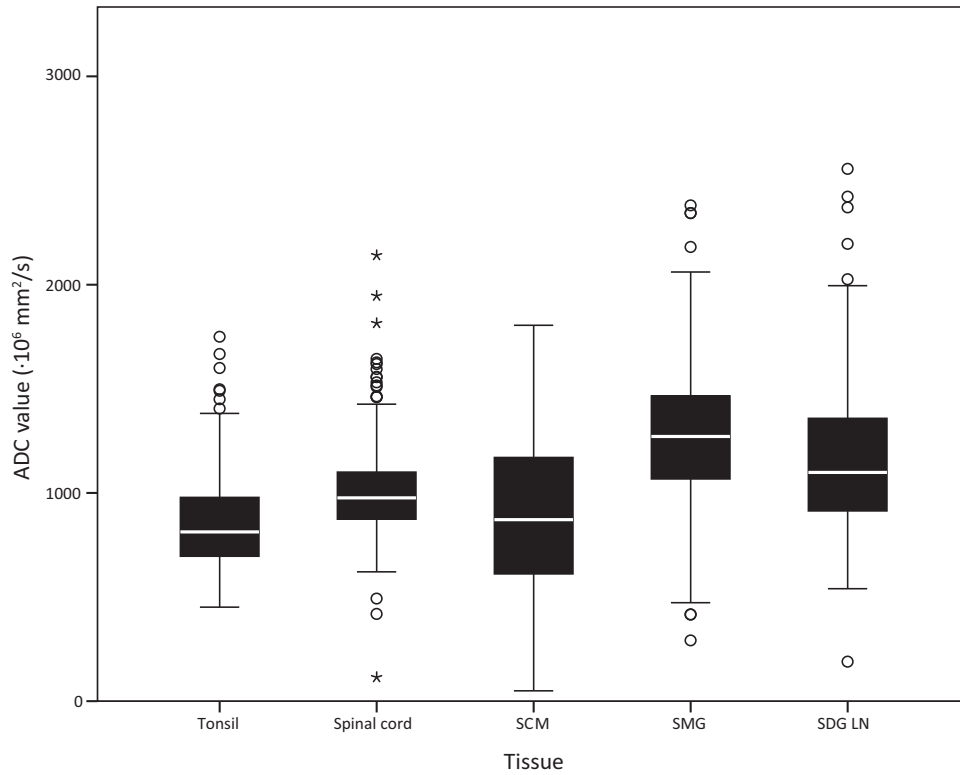
Before quantitative DWI can be applied in a multicenter study, knowledge is required about the reproducibility of ADC values within a subject, among different MR imaging systems, and among sequences.<sup>10</sup> This study is a first step to obtaining that knowledge.

In this study, we assessed the reproducibility of ADC values for different DWI sequences, MR imaging systems, and tissues in the head and neck. As expected, the variance in ADC values per subject per tissue is the smallest if the subject is measured on the same MR imaging system with the same sequence. The EPI-DWI-6b sequence showed the best reproducibility for all compared tissues, though this sequence was not available on all MR imaging systems. The EPI-DWI-2b sequence had a slightly lower reproducibility than the EPI-DWI-6b. Advantages of EPI-DWI-2b are a shorter acquisition time and being more widely clinically available. ADC measurements in the spinal cord and tonsil were the most precise and reproducible. Because the spinal cord is almost always present in the FOV during a head and neck study, this tissue can potentially be used as a reference. It also has the advantage of being rarely affected by malignancy; this advantage is in

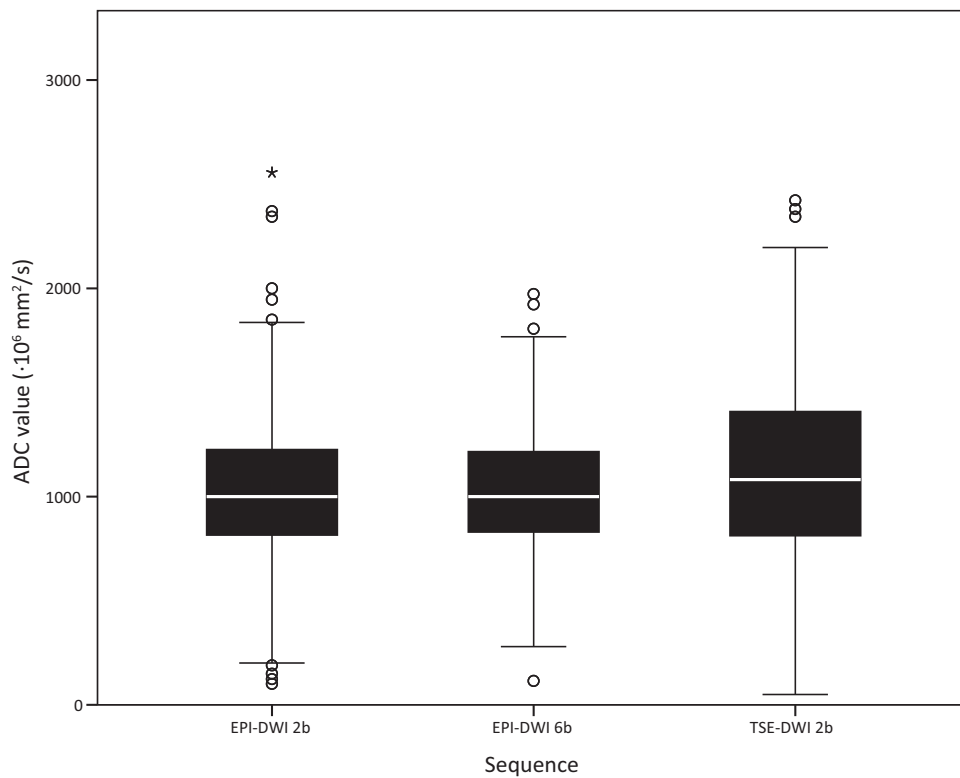
**Table 4: Actual ADC values ( $10^{-6} \text{ mm}^2/\text{s}$ ) and standard error of ADC measurement ( $10^{-6} \text{ mm}^2/\text{s}$ ) for all subjects and MRI systems**

|                    | EPI-DWI-2b       |       | EPI-DWI-6b       |       | TSE-DWI-2b       |       | Total per Tissue |       |
|--------------------|------------------|-------|------------------|-------|------------------|-------|------------------|-------|
|                    | Median (IQR)     | SEM   | Median (IQR)     | SEM   | Median (IQR)     | SEM   | Median (IQR)     | SEM   |
| Tonsil             | 791 (675–876)    | 134.2 | 746 (674–857)    | 119.6 | 1089 (839–1272)  | 203.0 | 813 (694–980)    | 151.2 |
| Spinal cord        | 950 (868–1053)   | 194.4 | 950 (865–1016)   | 170.6 | 1076 (908–1303)  | 204.2 | 976 (873–1100)   | 190.1 |
| SCM                | 990 (782–1276)   | 221.6 | 1084 (810–1317)  | 210.5 | 534 (286–822)    | 285.0 | 872 (611–1171)   | 237.8 |
| SMG                | 1257 (1090–1462) | 247.0 | 1233 (1066–1362) | 222.5 | 1392 (1030–1638) | 431.2 | 1271 (1066–1468) | 295.5 |
| SDG LN             | 1042 (809–1211)  | 307.9 | 1027 (870–1174)  | 242.9 | 1393 (1124–1709) | 322.0 | 1099 (910–1360)  | 291.0 |
| Total per sequence | 1000 (815–1226)  | 216.6 | 1000 (830–1217)  | 190.3 | 1082 (812–1414)  | 284.5 | 1020 (819–1273)  | 238.3 |

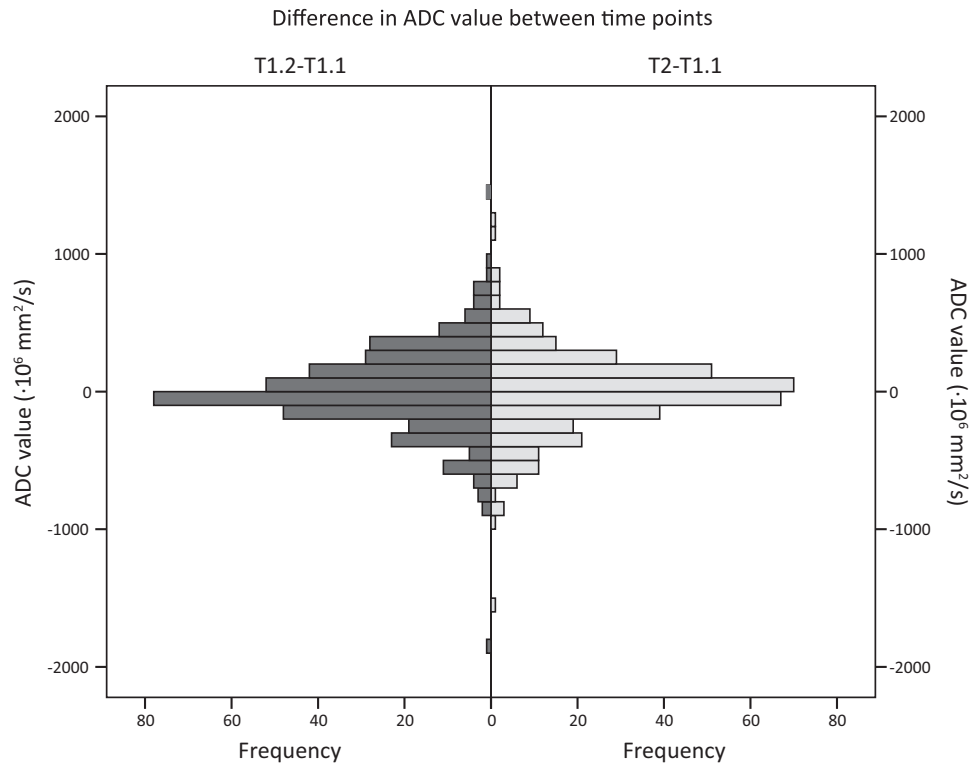
**Note:**—SCM indicates sternocleidomastoid muscle; SMG, submandibular gland; SDG LN, subdigastric lymph node.



**FIG 3.** Boxplots showing the distribution of ADC values ( $\times 10^{-6} \text{ mm}^2/\text{s}$ ) per sequence. The *points* are outliers (ie,  $>1.5$  IQRs away from the 25th or 75th percentile). The *asterisk* is an extreme outlier (ie,  $>3$  IQRs away from the 25th or 75th percentile).



**FIG 4.** Boxplots showing the distribution of ADC values ( $\times 10^{-6} \text{ mm}^2/\text{s}$ ) per tissue. The *points* are outliers (ie,  $>1.5$  IQRs away from the 25th or 75th percentile). The *asterisks* are extreme outliers (ie,  $>3$  IQRs away from the 25th or 75th percentile).



**FIG 5.** Histograms showing the difference in ADC values ( $\times 10^{-6}$  mm<sup>2</sup>/s<sup>2</sup>) compared with the first scan (T1.1). T1.1 and T1.2 were obtained on the same day. T2 was 1 month later. T1.2–T1.1: mean,  $6 \times 10^{-6}$  mm<sup>2</sup>/s<sup>2</sup>; SD,  $310 \times 10^{-6}$  mm<sup>2</sup>/s<sup>2</sup>. T2–T1.1: mean,  $-5 \times 10^{-6}$  mm<sup>2</sup>/s<sup>2</sup>; SD,  $310 \times 10^{-6}$  mm<sup>2</sup>/s<sup>2</sup>.

contrast to the tonsils, which are absent in case of tonsillectomy and frequently prove to be the location of an initially unknown primary tumor.<sup>17</sup> Therefore, the spinal cord seems to be the most suitable to serve as reference tissue.

DWI is frequently used in oncologic imaging.<sup>18,19</sup> Previous studies have shown the potential of DWI in diagnosing malignancies in the head and neck area, response prediction, and differentiation between treatment-induced tissue changes and residual or recurrent disease.<sup>6,20,21</sup> However, these studies were conducted in a single institution, without variance in MR imaging systems and protocol. Quantitative MR imaging parameters (eg, ADC) can differ substantially among MR imaging systems and imaging protocols.<sup>22</sup> This difference was also confirmed in the present study. We performed 3 examinations on 5 MR imaging systems on healthy subjects. This study validates differences in ADC values being statistically significant for sequences, MR imaging systems, and the interaction between MR imaging systems and sequences.

Verhappen et al<sup>8</sup> found TSE-DWI to be more reproducible among observers than EPI-DWI in a single-center, single-system study on primary tumors and lymph nodes of 12 patients with head and neck cancer. In the current multicenter, multisystem study, ADC values derived with the EPI-DWI-6b sequence were the most reproducible in healthy subjects with time, followed by EPI-DWI-2b. TSE-DWI-2b was the least reproducible sequence. These different findings may be attributed to the included subjects: healthy volunteers in the current study and patients with head and neck malignancies, which had diffusion restrictions in the study by Verhappen et al.<sup>8</sup> TSE-DWI-2b has inherently lower SNR,<sup>23</sup> which limits the reproducibility in healthy tissue, whereas it does not have geometric distortion and is apparently sensitive

enough to detect diffusion restriction. In the current study, ROIs were drawn on  $b=0$  s/mm<sup>2</sup> images in visual correlation with anatomic T2 images. Because EPI-DWI has a higher SNR, small structures (eg, benign lymph nodes) are more easily visualized. EPI-DWI may therefore be more appropriate for the evaluation of small structures. In a study by Vandecaveye et al,<sup>20</sup> 57% of malignant lymph nodes had a diameter of  $<1$  cm; therefore, appropriate evaluation of small (apparently benign) structures is vital. Verhappen et al<sup>8</sup> drew ROIs on ADC maps of malignant tissue that showed diffusion restriction. Especially, DWI of primary tumors in the head and neck area may have geometric distortion due to the tumor location at the air-tissue interface. In that case, geometric distortion of EPI techniques may reduce reproducibility among observers.<sup>8</sup>

There is also a difference in reproducibility among various tissues in the head and neck area. On all MR imaging systems and sequences, ADC values of the submandibular gland were the least precise (Table 4). An explanation for the relatively poor reproducibility might be the intrinsic physiologic changes in salivary glands during the time of day. ADC values in subdiaphragmatic lymph nodes have a relatively poor reproducibility (Table 4). Subdiaphragmatic lymph nodes are often too small for drawing reliable ROIs, particularly in healthy subjects. Moreover, lymph nodes are prone to changes with time (eg, due to frequently occurring inflammation in the head and neck area). In contrast, ADC values of the spinal cord and the tonsil are the most reproducible within subjects. In 87% of the images, a region of interest could be drawn on the tonsils; this percentage was lower than for the other tissues (range, 97%–98%) (Table 2). In healthy subjects, the tonsils are sometimes too small for reliably drawing a region of interest on

DWI. However, if the tonsils are large enough to allow the assessment of ADC values, these values appear to be relatively stable with time within a subject; this stability results in the relatively high precision and reproducibility of ADC measurements. The sternocleidomastoid muscle has intermediate reproducibility. Small changes in ADC values of muscle tissue may be explained by small differences in muscle tone with time.

Sasaki et al<sup>10</sup> previously assessed the reproducibility of ADC measurements in the brain among MR imaging systems, imaging protocols on different time points, and different institutions. It was concluded that there was significant variability in ADC values depending on the coil systems, imagers, vendors, and field strengths. However, only 3 of 10 patients were imaged more than once on the same MR imaging system. In our study, all patients were imaged multiple times on the same MR imaging system, in different institutions, and with a time interval of at least a month between imaging. We found significant differences between MR imaging systems and sequences.

The present study shows that though the physiology of healthy subjects may change with time, ADC values obtained within 1 person and with the same MR imaging system, protocol, and sequences immediately after the first scan and with an interval of at least 1 month have a low variance (ie, the intrasubject variance is small) (Fig 5). This finding indicates that ADC measurements are reproducible and independent of time. The spinal cord and tonsil are the tissues with the lowest ADC variability when different MR imaging systems, protocols, and sequences are used.

This study had some limitations. We included only healthy subjects with a broad age range for whom a stable physiologic status with time for all normal tissues can only be assumed. On the basis of Fig 5, the influence of time appears to be limited, with mean ADC differences being less than  $10 \times 10^{-6} \text{ mm}^2/\text{s}^2$  among measurements. The stability of the MR imaging systems and sequences used also needs to be assumed. Furthermore, the study population was too small to calculate a conversion factor for different MR imaging systems. A group size of  $\geq 50$  subjects is needed to calculate such a conversion factor.<sup>13</sup>

## CONCLUSIONS

The smallest range of ADC values can be obtained by imaging a subject on the same MR imaging system with an EPI-DWI with 6 b-values. Of the investigated tissues, the spinal cord shows the least variance and therefore should serve as reference tissue in the head and neck region.

## REFERENCES

- Jemal A, Bray F, Center MM, et al. **Global cancer statistics.** *CA Cancer J Clin* 2011;61:69–90
- de Bree R, Castelijns JA, Hoekstra OS, et al. **Advances in imaging in the work-up of head and neck cancer patients.** *Oral Oncol* 2009;45:930–35
- Le Bihan D, Breton E, Lallemand D, et al. **Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging.** *Radiology* 1988;168:497–505
- Chenevert TL, Meyer CR, Moffat BA, et al. **Diffusion MRI: a new strategy for assessment of cancer therapeutic efficacy.** *Mol Imaging* 2002;1:336–43
- Thoeny HC, De Keyzer F, King AD. **Diffusion-weighted MR imaging in the head and neck.** *Radiology* 2012;263:19–32
- Vandecaveye V, Dirix P, De Keyzer F, et al. **Predictive value of diffusion-weighted magnetic resonance imaging during chemoradiotherapy for head and neck squamous cell carcinoma.** *Eur Radiol* 2010;20:1703–14
- Kim S, Loevner L, Quon H, et al. **Diffusion-weighted magnetic resonance imaging for predicting and detecting early response to chemoradiation therapy of squamous cell carcinomas of the head and neck.** *Clin Cancer Res* 2009;15:986–94
- Verhappen MH, Pouwels PJ, Ljumanovic R, et al. **Diffusion-weighted MR imaging in head and neck cancer: comparison between half-Fourier acquired single-shot turbo spin-echo and EPI techniques.** *AJNR Am J Neuroradiol* 2012;23:1239–46
- Padhani AR, Liu G, Koh DM, et al. **Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations.** *Neoplasia* 2009;11:102–25
- Sasaki M, Yamada K, Watanabe Y, et al. **Variability in absolute apparent diffusion coefficient values across different platforms may be substantial: a multivendor, multi-institutional comparison study.** *Radiology* 2008;249:624–30
- Do RK, Chandarana H, Felker E, et al. **Diagnosis of liver fibrosis and cirrhosis with diffusion-weighted imaging: value of normalized apparent diffusion coefficient using the spleen as reference organ.** *AJR Am J Roentgenol* 2010;195:671–76
- Burdette JH, Elster AD, Ricci PE. **Calculation of apparent diffusion coefficients (ADCs) in brain using two-point and six-point methods.** *J Comput Assist Tomogr* 1998;22:792–94
- Carstensen B. *Comparing Clinical Measurement Methods: A Practical Guide.* Hoboken: John Wiley & Sons; 2010:59–67
- Carstensen B, Simpson J, Gurrin LC. **Statistical models for assessing agreement in method comparison studies with replicate measurements.** *Int J Biostat* 2008;4:Article 16
- Akaike H. **A new look at the statistical model identification.** *IEEE Trans Automat Contr* 1974;19:716–23
- Bryman, A, Cramer, D. *Quantitative Data Analysis with SPSS Release 12 and 13: A Guide for Social Scientist.* New York: Routledge; 2005:147–77
- Cianchetti M, Mancuso AA, Amdur RJ, et al. **Diagnostic evaluation of squamous cell carcinoma metastatic to cervical lymph nodes from an unknown head and neck primary site.** *Laryngoscope* 2009;119:2348–54
- Koh DM, Padhani AR. **Diffusion-weighted MRI: a new functional clinical technique for tumour imaging.** *Br J Radiol* 2006;79:633–35
- Sinkus R, Van Beers BE, Vilgrain V, et al. **Apparent diffusion coefficient from magnetic resonance imaging as a biomarker in oncology drug development.** *Eur J Cancer* 2012;48:425–31
- Vandecaveye V, De Keyzer F, Vander Poorten V, et al. **Head and neck squamous cell carcinoma: value of diffusion-weighted MR imaging for nodal staging.** *Radiology* 2009;251:134–46
- Vandecaveye V, De Keyzer F, Nuyts S, et al. **Detection of head and neck squamous cell carcinoma with diffusion weighted MRI after (chemo)radiotherapy: correlation between radiologic and histopathologic findings.** *Int J Radiat Oncol Biol Phys* 2007;67:960–71
- Pagani E, Hirsch JG, Pouwels PJ. **Intercenter differences in diffusion tensor MRI acquisition.** *J Magn Reson Imaging* 2010;31:1458–68
- Reiser M, Faber SC. **Recent and future advances in high-speed imaging.** *Eur Radiol* 1997;7:166–73