FULL PAPER

Magnetic Resonance in Medicine

An optimal acquisition and post-processing pipeline for hybrid IVIM-DKI in head and neck

Nienke D. Sijtsema[1,2] iD | **Steven F. Petit**[1] | **Dirk H. J. Poot**[2,3] iD | **Gerda M. Verduijn**[1] | **Aad van der Lugt**[2] | **Mischa S. Hoogeman**[1,4] | **Juan A. Hernandez-Tamames**[2]

[1]Department of Radiation Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands

[2]Department of Radiology and Nuclear Medicine, Erasmus MC, Rotterdam, The Netherlands

[3]Department of Medical Informatics, Erasmus MC, Rotterdam, The Netherlands

[4]Department of Medical Physics & Informatics, HollandPTC, Delft, The Netherlands

Correspondence
Nienke Sijtsema, Department of Radiation Oncology, Erasmus MC Cancer Institute, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands.
Email: n.sijtsema@erasmusmc.nl

Funding information
This work was funded by a research grant from Elekta AB (Stockholm, Sweden). SP would like to acknowledge financial support from The Dutch Cancer Society project number 12141.

Purpose: To optimize the diffusion-weighting b values and postprocessing pipeline for hybrid intravoxel incoherent motion diffusion kurtosis imaging in the head and neck region.

Methods: Optimized diffusion-weighting b value sets ranging between 5 and 30 b values were constructed by optimizing the Cramér-Rao lower bound of the hybrid intravoxel incoherent motion diffusion kurtosis imaging model. With this model, the perfusion fraction, pseudodiffusion coefficient, diffusion coefficient, and kurtosis were estimated. Sixteen volunteers were scanned with a reference b value set and 3 repeats of the optimized sets, of which 1 with volunteers swallowing on purpose. The effects of (1) b value optimization and number of b values, (2) registration type (none vs. intervolume vs. intra- and intervolume registration), and (3) manual swallowing artifact rejection on the parameter precision were assessed.

Results: The SD was higher in the reference set for perfusion fraction, diffusion coefficient, and kurtosis by a factor of 1.7, 1.5, and 2.3 compared to the optimized set, respectively. A smaller SD (factor 0.7) was seen in pseudodiffusion coefficient. The sets containing 15, 20, and 30 b values had comparable repeatability in all parameters, except pseudodiffusion coefficient, for which set size 30 was worse. Equal repeatability for the registration approaches was seen in all parameters of interest. Swallowing artifact rejection removed the bias when present.

Conclusion: To achieve optimal hybrid intravoxel incoherent motion diffusion kurtosis imaging in the head and neck region, b value optimization and swallowing artifact image rejection are beneficial. The optimized set of 15 b values yielded the optimal protocol efficiency, with a precision comparable to larger b value sets and a 50% reduction in scan time.

KEYWORDS

b value optimization, diffusion-weighted imaging (DWI), head and neck, intravoxel incoherent motion diffusion kurtosis imaging (IVIM-DKI), non-Gaussian IVIM (NG-IVIM), swallowing artifact correction

1 | INTRODUCTION

In the last decade, the apparent diffusion coefficient has been shown to be a promising parameter for response assessment of head and neck cancer treated with (chemo)radiotherapy.[1,2] More recently, the benefit of obtaining additional parameters from DWI by intravoxel incoherent motion (IVIM) and diffusion kurtosis imaging (DKI) has been acknowledged.[3,4] Combining these models in hybrid IVIM-DKI allows for the simultaneous assessment of the diffusion coefficient (*D*), perfusion fraction (*f*), pseudodiffusion coefficient (*D**), and kurtosis (*K*). Both *f* and *D** are related to perfusion and could therefore provide insight in abnormal tumor perfusion. The latter parameter, *K*, is associated with restricted diffusion and cellularity,[5] which is commonly increased in tumors. However, full assessment of their potential as response biomarkers and, in the future, for application in radiotherapy treatment planning and treatment adaptation, requires robust and repeatable estimation of all IVIM-DKI parameters within a patient.[6]

IVIM-DKI parameter estimation is currently hampered by inefficient sampling of the DWI signal curve and motion artifacts. Inefficient sampling of the DWI signal curve, due to nonoptimal choice of *b* values, leads to unnecessary long scan times in order to achieve a similar precision as optimized sampling. Moreover, long scan times might increase the amount of motion corruption. Several efforts have been made to optimize *b* values for the monoexponential,[7-13] the IVIM model with direct[13-16] and segmented fitting,[17] and the kurtosis model[13] in a variety of tissues, although none of the mentioned papers address the head and neck region specifically. For complex models such as hybrid IVIM-DKI, optimizing *b* values is increasingly more difficult and, to the best of our knowledge, has not been done so far.

Additionally, the head and neck region is prone to several types of motion artifacts. Firstly, swallowing and coughing artifacts, which present themselves as severe signal dropout, could cause over- or underestimation of the DWI parameters. Current mitigation strategies mainly consist of specific patient instruction for not swallowing or coughing but are not always sufficient because both can be a reflex behavior. Secondly, head motion hampers parameter estimation by causing misalignment between scan volumes. This type of motion can be partially prevented using fixation of the patient, either in the form of padding, or in the case of radiotherapy-planning MRI, with an immobilization mask. Additionally, motion artifacts might be corrected after acquisition by registration; however, registration of high *b* value images can be problematic due to lower SNR.

Therefore, the goal of this study was to find an optimal acquisition and develop a postprocessing pipeline for hybrid IVIM-DKI DWI in the head and neck region. To this end, we optimized the *b* values of hybrid IVIM-DKI for the head and neck region, applied motion compensation, and investigated the effect of swallowing artifact rejection.

2 | METHODS

Acquisition optimization and postprocessing of the hybrid IVIM-DKI DWI data consisted of 3 stages. In the first stage, described in Section 2.1, the set of *b* values was optimized for different *b* value set sizes based on simulated ground truth parameter sets. Next, these *b* value sets were scanned in healthy volunteers, as described in Section 2.2. In the second stage, intra- and intervolume registration were applied to the acquired data as described in Section 2.3. The third stage consisted of swallowing artifact image rejection (Section 2.4). Lastly, Section 2.5 describes assessment of the parameter estimation precision of the pipeline. In this paper, the term postprocessing refers to all processing done after acquisition but before parameter estimation.

2.1 | Optimization of *b* values

The hybrid IVIM-DKI model is given by:

$$S_b(\boldsymbol{\theta}) = S_0 \left((1-f) \left(e^{-bD + \frac{1}{6}(bD)^2 K} \right) + f e^{-bD^*} \right), \quad (1)$$

where $S_b(\boldsymbol{\theta})$ is the signal at a specific *b* value (amount of diffusion weighting) as function of the parameters $\boldsymbol{\theta} \in \{S_0, f, D^*, D, K\}$, which are described in Table 1.

A ground truth set $\boldsymbol{\theta}$ was created using a Halton sequence[18] for the ground truth ranges shown in Table 1. A Halton sequence yields a sequence of low-discrepancy, pseudorandom numbers. These ranges were derived from the parameter values reported in several previous studies in head and neck tumors.[2,3,19-21] From these ground truth parameter sets, sets showing a *D** smaller than 6 times *D* were removed to be able to properly separate the 2 parameters. Some parameter combinations result in increased signal at high *b* values. To eliminate this nonphysical result, parameter sets in which the partial derivative of the model to the *b* value was larger than 0 at $b = 1500$ s/mm^2 were removed. After exclusion, 272 of 576 ground truth $\boldsymbol{\theta}$ were left.

The IVIM-DKI *b* value set was optimized for precision of parameter estimation within the available acquisition time with the cost function:

$$C(\boldsymbol{b}) = \sqrt{\overline{\left(relCRLB(\boldsymbol{\theta}, \boldsymbol{b})\right)^2}_{\theta,i}} \times T_{acq}(\boldsymbol{b}), \quad (2)$$

where *C* is the cost function value and $\boldsymbol{b}$ the set of *b* values; the bar with subscripts indicates mean over parameter *i* and $\boldsymbol{\theta}$, where *i* corresponds to *f*, *D**, *D*, and *K*.

TABLE 1 Overview of parameters in the hybrid IVIM-DKI model and minimum and maximum value chosen for the ground truth value ranges

Parameter	Description	Minimum	Maximum
S_0	Signal intensity at $b = 0$ s/mm^2	4275.8	7126.3
D	Diffusion coefficient	$0.25 \cdot 10^{-3}$ mm^2/s	$3.41 \cdot 10^{-3}$ mm^2/s
f	Perfusion fraction	0.09	0.42
$D*$	Pseudodiffusion coefficient/ apparent perfusion coefficient	$6.29 \cdot 10^{-3}$ mm^2/s	$237.39 \cdot 10^{-3}$ mm^2/s
K	Kurtosis	0.1	2.81

Note: S_0 in this table is the estimated signal intensity without T2 decay effects.

Abbreviation: IVIM-DKI, intravoxel incoherent motion diffusion kurtosis imaging.

In Equation (2) the relative Cramér-Rao lower bound (rel-CRLB) is

$$relCRLB\,(\boldsymbol{\theta}, \boldsymbol{b}) = \left(\boldsymbol{F}\,(\boldsymbol{\theta})^{-1}\right)_{i,i} / \theta_i^2, \tag{3}$$

with $\boldsymbol{F}\,(\boldsymbol{\theta})$, the Fisher information matrix:

$$F_{i,j}\,(\boldsymbol{\theta}) = -\mathbb{E}_S \left[\frac{\partial^2 \ln p(\boldsymbol{S}|\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right], \tag{4}$$

where and $p(\boldsymbol{S}|\boldsymbol{\theta})$ is the joint probability density function of the Rician-distributed measurements $\boldsymbol{S}$.

The CRLB, $\boldsymbol{F}\,(\boldsymbol{\theta})^{-1}$, gives a lower bound of the variance for the given parameter vector and is commonly used in experiment design.[22] Therefore, it is used as a measure of precision in this paper. The optimization of $C\,(\boldsymbol{b})$ was constrained to avoid negative b values and b values higher than $b = 1500$ s/mm^2. A maximum b value of 1500 s/mm^2 was chosen due to limited SNR in the head and neck region at higher b values in healthy tissue at 1.5 tesla. The acquisition time was given by $T_{acq} = (TE + RO) * n_{b-values} * n_{slices}$, with $TE = 61 + 0.015 * b_{max}$ and RO the readout time. This relation between the maximum b value and TE was experimentally obtained from the sequence implementation. The readout time RO was determined to be 120 ms for the scan protocol used in this paper.

Furthermore, a TE correction was applied to the signal S_0 of the ground truth voxels as the signal decreases when TE increases:

$$S_{0,corrected} = S_0 * e^{-\frac{TE}{T2}}. \tag{5}$$

The T2 was set to 80 ms. A maximum b value of 1500 s/mm^2 resulted in a modeled signal intensity at $b = 0$ s/mm^2 of 1500 to 2500. The noise level was set to 20, leading to an estimated intratumor SNR of 75-125 at $b = 0$ s/mm^2. Initial sets were chosen containing b values along the range of 0-1500 s/mm^2. These sets were optimized using the fminsearch algorithm in MatLab 2017a (MathWorks, Natick, MA). Optimization was carried out for a set size of 5, 6, 10,

15, 20, and 30 b values. The set of 20 b values was optimized first and initialized with approximately linearly distributed b values. Because the final b values were clustered, we subsequently restarted the optimization with varying initial distributions of b values over the 3 regimes (perfusion ($b = {\sim}0\text{-}200$ s/mm^2), free diffusion ($b = {\sim}200\text{-}800$ s/mm^2), and restricted diffusion $b > {\sim}800$ s/mm^2) to reduce the chance of ending in a local minimum. For each number of b values, the set with the lowest overall cost value was selected. The b value optimization code is available online at github.com/nsijtsema/IVIMDKI_b-value_optimization.

2.2 | MR scanning

Seventeen healthy volunteers (14 females, 3 males, mean age 26, age range 19-59) were imaged on a 1.5 T GE Optima MR450w GEM imaging system (General Electric Medical Systems, Waukesha, WI) with a 16-channel head and neck coil (General Electric Medical Systems, Waukesha, WI). The study was approved by the institutional review board (protocol 2014-096), and written informed consent was obtained from all volunteers. The imaging protocol consisted of a T2 periodically rotated overlapping parallel lines with enhanced reconstruction (scan time 5 minutes 34 seconds) followed by a single shot spin-echo EPI IVIM-DKI DWI acquisition (FOV 26×26 cm^2, 128×128 matrix, $2 \times 2 \times 4$ mm voxels, 0.2 mm interslice gap, TE = 81.8 ms, TR = 6700 ms, SENSE parallel imaging acceleration factor 2, number of averages = 1, 3 orthogonal diffusion directions) with a reference b value set (scan time 6 minutes 35 seconds) of the geometrical form $\left(0, 10a^1, \ldots, 10a^{n-1}\right)$ with $a = 1.3$ and the number of b values $n = 20$, similar to the approach used by Lu et al.[3] Next, 3 repeats of the optimized single shot spin-echo EPI IVIM-DKI DWI sets (scan time 11 minutes 17 seconds) were scanned with the same imaging parameters mentioned for the reference b value set, with exception of the b values. Volunteers were asked to lie still, except for the last optimized IVIM-DKI acquisition. During that acquisition, volunteers were asked to swallow 4 times so that most optimized b value set had at least 1 swallowing artifact. For all DWI

acquisitions, reversed readout gradient polarity for $b = 0$ s/mm^2 (scan time 1 minute 7 seconds) was collected for the purpose of distortion correction using reversed gradient polarity blips,[23,24] which were applied to all images. The odd and even slices of the DWI were acquired in interleaved fashion to minimize slice crosstalk. Total scan time of the protocol was 55 minutes.

To fit the acquisition in the available scan time, the optimized b value sets containing 5, 6, 10, 15, 20, and 30 b values were combined into 1 set. This was done by adding the b values of the longer subsequent set that were not already present in the shorter set to the end of the acquisition. The set of 5 was thus expanded with 2 b values not initially present in that set to be able to form the optimized set of 6. The set was then consecutively expanded to 10, 15, 20, and 30 b values in the same manner. Consequently, the b value sets were grouped in time to ensure the effect of motion is representative for a patient scan. Due to overlap between the b values in the optimized sets, the total b value set encompassed 34 b values. Table 2 contains the acquired b value sets.

2.3 | Registration

Two registration methods were compared with nonregistering: intervolume registration only and intra- and intervolume registration. The methods for both intra- and intervolume registration were obtained from Guyader et al.[25] In case of intravolume registration, the odd and even slices from each b value image were separated into 2 separate volumes. Subsequently, the 2 volumes were aligned by group-wise registration because the odd and even slices were acquired interleaved. Intervolume registration was carried out by pair-wise registration of each b value image to the $b = 0$ s/mm^2 image. Intravolume registration, if applicable, was performed before intervolume registration. All registrations were nonrigid and carried out with Elastix (version 4.9).[26] The parameter files from the registration approach[25] are available in the Elastix parameter file database. Registration was applied after distortion correction for all cases in the mentioned order. Intervolume registration errors were detected by manually assessing the imaging volumes and identifying the volumes with severe anatomical mismatches. To correct for registration errors, the transformations of the 2 well-registered b values that were closest in time were linearly interpolated and applied to the original image that contained the registration error.

Finally, the $b = 0$ s/mm^2 scan of each acquisition was registered to the reference $b = 0$ s/mm^2 acquisition using the same registration approach as used for the intervolume registration. The transformation of the $b = 0$ s/mm^2 was propagated to the remaining images of the acquisition.

2.4 | Swallowing artifact image rejection

Swallowing artifacts presented as severe signal dropout in the individual b value images were detected by visual inspection and subsequently rejected. Because the data were acquired using 3 sequential orthogonal gradient directions, artifacts present in 1 diffusion direction resulted in rejection of all 3 directions to maintain isotropic weighting in the fit.

2.5 | Assessment

Regions of interest (ROIs) were drawn inside both tonsils of each volunteer based on the first acquired $b = 0$ s/mm^2 volume. Subsequently, the ROIs were propagated to the other DWI images in the scanned series. For analysis, the tonsils were regarded as a single organ. Voxel-wise fitting was performed with an in-house fitting script in MatLab 2017a (MathWorks), which employs a variable starting point method before direct fitting of the biexponential with the most suitable starting point. The range in which starting points were chosen was the same as the ranges chosen for the optimization of f, D^*, and D. For S_0, the range was widened to 200 to 5000, and for K the upper bound was rounded to 3. Fitting constraints were set to prevent severe outliers but aiming to avoid skewing the distribution at the edge of the physiologically plausible parameter values. The used constraints were S_0 in [0, 10 000] [arbitrary units], f in [−1, 1] [–], D^* in [0, 0.8] mm^2/s, D in [0, 0.02] mm^2/s, and K in [−5, 5] [–] as a compromise between unconstrained and more strictly constrained fitting.

2.5.1 | Comparison of the optimized set of 20 b values with the reference set

To assess the change in precision due to b value optimization, the nonoptimized reference set was compared with the optimized set of 20 b values from the first optimized acquisition

	b Values
Reference	0, 10, 1460, 20, 1120, 20, 870, 30, 670, 40, 510, 50, 390, 60, 300, 80, 230, 110, 180, 140
Optimized	0, 20, 780, 1500, 130, 790, 640, 80, 1500, 570, 770, 770, 80, 1500, 780, 1500, 10, 790, 1500, 790, 1500, 80, 750, 1500, 80, 760, 790, 80, 750, 280, 1500, 80, 790, 10

TABLE 2 Overview of the order in which the b values were acquired for the reference acquisition and the combined optimized IVIM-DKI acquisition

Abbreviation: IVIM-DKI, intravoxel incoherent motion diffusion kurtosis imaging.

in terms of mean and SD of the 4 parameters of interest (f, $D*$, D, K) within the ROI. The SD in the ROI consists of both underlying physiological differences and noise. Because the underlying physiological differences are constant, the SD within the ROI was used as a measure of precision. Only the fully corrected data (with applied distortion correction, intra- and intervolume registration) were used in this comparison. A Wilcoxon signed-rank test was used to test significance for both metrics and all parameters, leading to a total of 8 tests.

2.5.2 | Repeatability assessment for type of registration and set size

To assess the repeatability between the 2 identical optimized nonswallowing acquisitions for the different set sizes and registration scenarios, the coefficient of repeatability (CR)[27] was used. The third repetition contained swallowing artifacts and was therefore not used in this assessment.

$$CR = 2.77 \sqrt{\frac{1}{2n} \sum_j \left(\overline{p_{j,A}} - \overline{p_{j,B}} \right)^2}, \quad (6)$$

where $\overline{p_{j,k}}$ is the ROI mean of the parameter of interest for each subject j at time point k (either A or B), and n is the number of subjects. The workflow for obtaining the CR is schematically depicted in Figure 1. Then, the relative CR (relCR) is defined as $relCR = \frac{CR}{\overline{P}} * 100$, for which $\overline{P}$ is the overall ROI mean over the 2 acquisitions across all volunteers: $\overline{P} = \frac{1}{2n} \sum_{j,k} \overline{p_{j,k}}$. Similarly, the relative difference (relDifference) in parameter mean is defined as $relDifference = \frac{\overline{p_{j,A}} - \overline{p_{j,B}}}{\overline{P}} * 100$.

The optimized set of 30 b values was compared to the sets of 20, 15, 10, 6, and 5 in terms of relCR and relDifference. For set size assessment, only the fully corrected data (applied distortion correction, intra- and intervolume registration) were used. Wilcoxon signed-rank tests were applied to test if relDifference was statistically significantly different from 0 for all set sizes and parameters, resulting in 24 tests.

For comparison of registration methods, only the optimized set size of 30 b values was used. Comparisons in relDifference and relCR were carried out between the no registration and intervolume-only registration cases as well as between the no registration and intra- and intervolume registration cases. Wilcoxon signed-rank tests were applied to test if the relDifference was different from 0 for all registration cases and all parameters, resulting in 12 tests.

2.5.3 | Assessment of swallowing artifact rejection

In case of swallowing artifact image rejection, the ROI mean before rejection was compared to the ROI mean after rejection, and a Wilcoxon signed-rank test was applied. Comparisons were done for each set size applying intra- and intervolume registration, resulting in 24 comparisons. Additionally, in case a significant difference ($P \leq .05$) was found before and after swallowing artifact rejection, the mean before and after rejection was compared to that of the second optimized acquisition without swallowing to assess improvement in accuracy, leading to another 18 comparisons.

3 | RESULTS

Figure 2 shows an example of the tonsil ROIs and corresponding parametric maps. One volunteer was excluded because the tonsils of this volunteer had been removed.

3.1 | Optimization of b values

Table 3 shows the resulting optimized sets for 5, 6, 10, 15, 20, and 30 b values. Note that all b values are rounded to multiples of 10 to comply with the scanner's requirements. The increase in cost function value due to rounding was less than 1%, except for the set of 5 b values where it was 1.8%. The mean and SD over all volunteers and set sizes for the 2 repeated fully registered nonswallowing acquisitions of the mean f, $D*$, D, and K within the ROIs were 0.12 ± 0.06 [–] for f, 0.07 ± 0.03 mm^2/s for $D*$, $0.8 \cdot 10^{-3} \pm 0.2 \cdot 10^{-3}$ mm^2/s for D, and 0.73 ± 0.53 [–] for K. The obtained values of S_0 were 1021 ± 136 [arbitrary units.].

Comparing the relative CRLB from the reference set with the optimized set in the simulated ground truth parameter sets, demonstrated that the reference set was expected to have a factor 2.2 higher variance than the optimized set for K. Slightly lower variance was expected in the reference set for f, D, and $D*$ with factors 0.5, 0.59, and 0.56, respectively, compared to the optimized set.

3.1.1 | Comparison of the optimized set of 20 b values with the reference set

The SD for D, K, and f was significantly lower in the optimized set of 20 b values compared to the reference set that also had 20 b values (D: $P = .04$, K: $P = .003$, f: $P = .002$), as is shown in Figure 3. This corresponded to a 1.7, 1.5, and 2.3 times higher SD in the reference set in comparison to the optimized set in f, D, and K, respectively. The SD of $D*$ was significantly higher in the optimized set ($P = .006$), corresponding with a 0.7 lower SD in the reference set versus the optimized set. On average over all parameters, the improvement in SD was a factor 1.55. A statistically significant difference in mean was only seen for $D*$ ($P = .01$). A nonphysiological mean for K ($K < 0$) was found in 3 volunteers
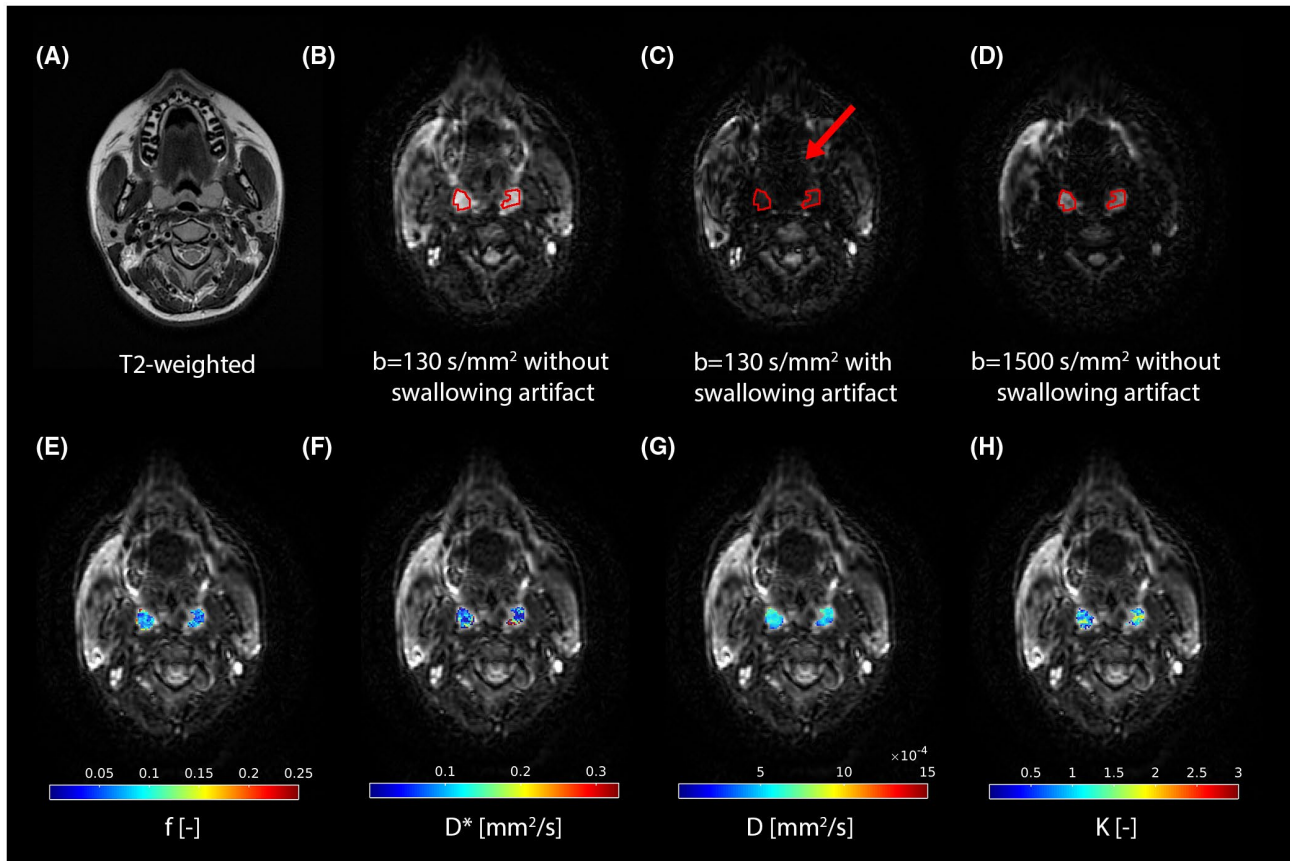
FIGURE 1 Overview of the workflow for obtaining the difference in mean and CR. Each acquisition (denoted A and B) contained multiple *b* values. After distortion correction and registration, a fit can be made for each voxel, yielding estimates of the 4 IVIM-DKI parameters. Subsequently, the mean in the tonsils was calculated over both tonsils for each parameter for acquisitions A and B, and the mean of acquisition A was subtracted from the mean of acquisition B to yield d_j. These steps were repeated for each volunteer. Finally, the average difference in mean and CR were calculated. CR, coefficient of repeatability; IVIM-DKI, intravoxel incoherent motion diffusion kurtosis imaging

for the reference set and in 1 volunteer for the optimized set of 20 *b* values. Additionally, a nonphysiological mean for *f* ($f < 0$) was found in 1 volunteer in the reference set and in none of the volunteers in the optimized set of 20 *b* values. Boxplots of the difference in mean of the optimized set versus the reference set can be found in Supporting Information Figure S1.

3.1.2 | The effect of the number of *b* values on parameter precision

Figure 4 shows relDifference and relCR each for set size. The relCR of *D* was 37.7% for a set size of 30 and varied only slightly for set sizes of 20 and 15. However, for set size smaller than 15 *b* values, a considerable increase in relCR

was observed up to 73.7% for a set size of 6. The pattern in *D** was similar to *D*, with relCR ranging from 74.3% in the set size of 20 *b* values up to 110.4% in the set size of 5 *b* values. However, the relCR of *D** in set size 30 was 124.2%, and therefore notably higher than any of the other set sizes. For *f*, the relCR was lowest in set size 15 and 20 and slightly higher in the other set sizes. A slight decrease from 116.4% to 97.7% in relCR was seen in set size 20 compared to 30 in *K*. When removing more *b* values from the set, relCR of *K* increased continuously up to 292.8% in the set of 5 *b* values. Significant differences in parameter mean were found in the set size 10 for *f* (31.6%) and the set size 20 for *D** (23.0%). However, the relDifference of *f* was substantial with a range of 17.1% to 31.6% in all set sizes. Boxplots of the bias of set size 30 versus the other set sizes can be found in Supporting Information Figure S2.

FIGURE 2 (A) Axial T2-weighted slice with (B) corresponding $b = 130$ s/mm^2 axial slice with the tonsils outlined in red and (C) identical slice to B but affected by a swallowing artifact. In (D), the same slice is shown but for $b = 1500$ s/mm^2. In the bottom row, parametric maps for the tonsils are shown as color overlays on the corresponding $b = 0$ s/mm^2 slice with (E) f, (F) D^* in mm^2/s, (G) D in mm^2/s, and (H) K. D, diffusion coefficient, D^*, pseudodiffusion coefficient; f, perfusion fraction; K, kurtosis

TABLE 3 Rounded results of the optimization

	b **Values [s/mm^2]**
5 b values	0, 20, 130, 780, 1500
6 b values	0, 20, 80, 640, 790, 1500
10 b values	0, 20, 2×80, 570, 2×770, 780, 2×1500
15 b values	0, 10, 2×80, 130, 570, 2×770, 2×780, 790, 4×1500
20 b values	0, 10, 3×80, 130, 570, 2×770, 2×780, 3×790, 6×1500
30 b values	0, 2×10, 6×80, 280, 2×750, 760, 2×770, 2×780, 5×790, 8×1500

3.2 | Registration

When only intervolume registration was applied, correction of severe registration errors was necessary. No registration errors were found in case intravolume registration was also applied. Figure 5 shows relDifference and relCR for each registration scenario. Intervolume-only registration showed a smaller relCR by 25% to 40% in f compared to the other scenarios. A similar effect was seen in D, where the decrease was around 10%. However, the opposite effect was seen in D^* and K, where relCR increases of 1% to 23% and 30% to

42% were observed, respectively. Significant differences in parameter mean were found for no registration (31.5%) and intervolume-only registration (22.2%) for f, leaving the full registration as the only scenario without significant difference in parameter mean in any parameter. Nevertheless, the 32.5% relDifference for f in the full registration scenario was still substantial. Boxplots of the difference in mean of the fully registered set versus the other 2 registration procedures can be found in Supporting Information Figure S3. Parameter maps for each registration scenario and parameter are shown for 1 volunteer in Supporting Information Figure S4.
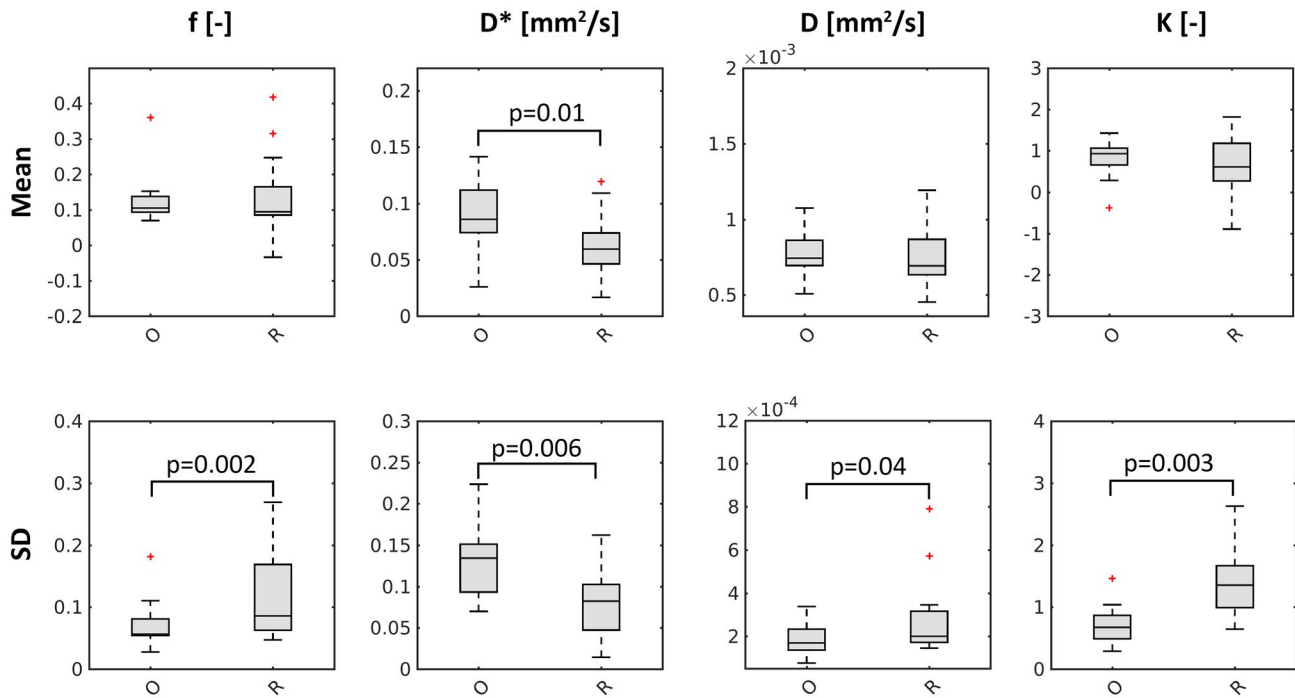
FIGURE 3 Boxplots showing the distribution of the mean and SD in the tonsils over the volunteers for each of the 4 parameters of interest: f, D^*, D, and K. The optimized set of 20 b values is marked "O" and the reference set "R." P values are indicated in case $P \leq .05$ according to the Wilcoxon signed rank test
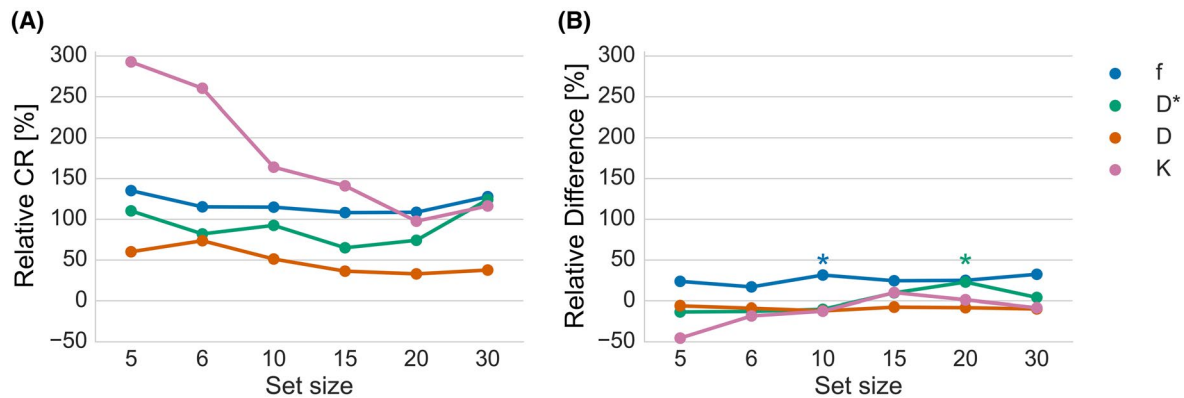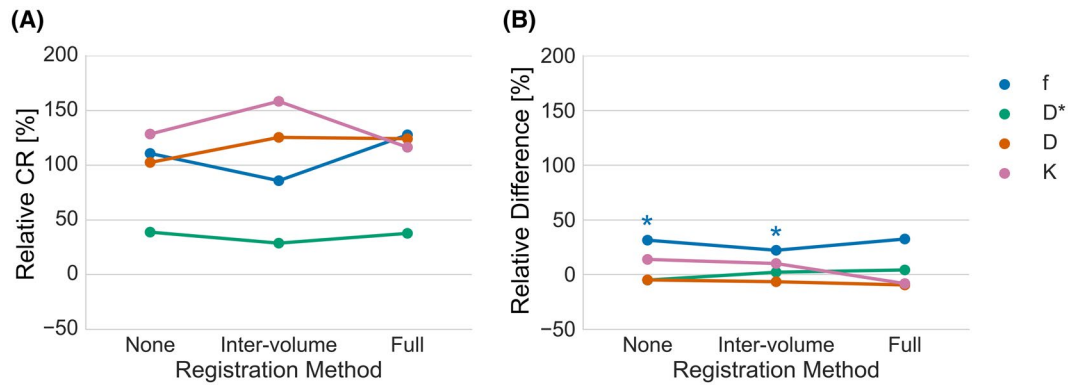

FIGURE 4 (A) Relative CR and (B) relative difference in mean between 2 identical acquisitions shown as a percentage of the overall mean of the 2 acquisitions for the 4 parameters of interest: f, D^*, D, and K. Significant difference in mean based on Wilcoxon signed rank test was found in f for set size 10 and D^* for set size 20. These points are denoted by an asterisk marker. For an explanation of the metrics, see Figure 1

3.3 | Swallowing artifact rejection

Table 4 shows the average percentage of rejected b value volumes per set size due to swallowing artifact presence. Significant differences in ROI mean between artifact-rejected and nonartifact-rejected data were found for a set size of 5 for all parameters (f: $P = .001$, D^*: $P = .007$, D: $P = .0005$, K: $P = .003$), as well as in set sizes 10 ($P = .01$), 15 ($P = .008$), and 20 ($P = .008$) for f and set size 15 ($P = .006$) and 20 ($P = .03$) for D. In all mentioned cases, except D, a significant difference in ROI mean was observed between the nonrejected data and the second optimized acquisition. This difference was resolved after artifact image rejection, as is

illustrated by Figure 6 for the mean perfusion fraction, indicating repeatability is improved after artifact image rejection. Boxplots for the other parameters can be found in Supporting Information Figure S5. Supporting Information Figure S6 contains scatter plot of the data before and after swallowing rejection plotted versus the baseline for all parameters and set sizes.

4 | DISCUSSION

Hybrid IVIM-DKI is promising for response assessment of head and neck cancer during (chemo)radiotherapy.[3,4]

(A)

(B)

FIGURE 5 (A) Relative CR and (B) relative difference in mean between 2 identical acquisitions shown as a percentage of the overall mean of the 2 acquisitions for the 4 parameters of interest: f, D^*, D, and K. Full registration comprises intra- and intervolume registration. Significant difference in mean based on Wilcoxon signed rank test was found in f for no registration and intervolume-only registration. These points are denoted by an asterisk. For an explanation of the metrics, see Figure 1

TABLE 4 Average percentage of rejected data per set size over 16 volunteers

Set size	Average amount rejected b values (%)
30	9.6
20	15.3
15	15.0
10	12.5
6	4.2
5	16.3

However, inefficient sampling of b values and motion corruption hamper precise parameter estimation in hybrid IVIM-DKI and thus diminish its potential for response assessment. In this paper, we showed that acquisition optimization and motion correction improves parameter estimation precision of the parameters of interest. Additionally, parameter repeatability metrics did not improve for sets larger than 15 b values, suggesting that 15 b values is sufficient. Possibly because image acquisition noise is not the dominant error term for larger sets. No benefit of registration was found in our data, but swallowing artifact image rejection was beneficial, especially if a larger part ($>10\%$) of the data is affected by such artifacts. To the best of our knowledge, this is the first study that addressed combined acquisition optimization and motion correction for hybrid IVIM-DKI in head and neck.

Most b values in the optimal sets were near or above 800 s/mm^2 (restricted diffusion regime), followed by b values below 200 s/mm^2 (perfusion regime), whereas relatively little b values were chosen in between in the free diffusion regime. Our findings suggest that the distribution of the b values reflects the uncertainty in the main parameters that rely on each regime. D was the most reliable parameter, followed by f and

D^*. K relies mainly on high b values and was the least reliable parameter.

Using a higher maximum b value than 1500 s/mm^2 is expected to improve the precision of K as long as the signal intensity at the maximum b value is well above the noise floor to avoid parameter estimation bias. Since the model predicts an increase in signal at very high b values, it is not expected that the optimization reaches a natural maximum b value. Instead, it would venture into b values where the kurtosis model is not valid. In this work, we chose a maximum b value of $b = 1500$ s/mm^2 due to the limited SNR at this b value in the head and neck region at 1.5 tesla to avoid the noise floor and the related parameter estimation bias.

Previous studies on b value optimization were carried out for the monoexponential DWI model[7-9] and the IVIM model.[14,15,17] These studies have largely reported an amount of unique b values equal to the minimum of b values needed to estimate all parameters. In this work, however, sets larger than 5 b values always consisted of more than 5 unique b values. Nonetheless, in the monoexponential model Brihuega-Moreno et al[9] also reported more variation in b values than the minimum needed for estimation; however, this was in the presence of large parameter ranges and without maximum b value constraint. This could be in line with this study because the parameter ranges applied in this study were quite broad, and due to inclusion of K, the b value range was also larger. Additionally, a larger distribution was seen in more complex problems, for example, in optimizing gradient strengths and directions for diffusion kurtosis imaging.[28] Yet, even then discretization is seen.

The optimization technique discussed in this paper consists of a general framework and can be tailored to specific interests by weighting the parameters or even choosing another model. We chose an equal weighting for all the parameters of interest (f, D^*, D, and K). However, if there is no
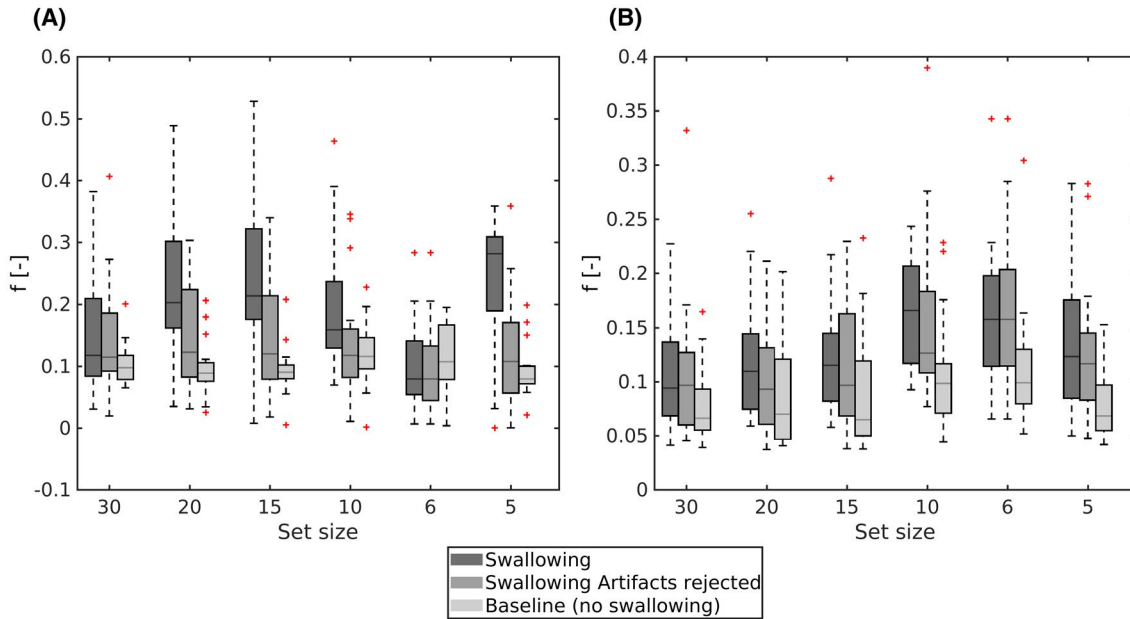
FIGURE 6 Boxplot of the mean (A) and SD (B) of *f* over 16 volunteers for each set size of 5 to 30 *b* values. The acquisition including swallowing artifacts (darkest gray) performs worse than the same acquisition after swallowing artifact correction (middle gray) when compared to the baseline acquisition (lightest gray), which did not contain any swallowing artifacts

interest in 1 specific parameter, it can be left out of the optimization. Similarly, if 1 parameter is more important than the others, a higher weight can be assigned to this parameter. Therefore, specifying the objective function for *b* value optimization should be adapted according to specific interests and is therefore subjective in nature.[17]

Standardization of DWI across imaging systems is challenging.[29] The presented optimization framework depends on 2 parameters that are system-dependent: SNR and TE. The *b* value set is quite robust to SNR, with only about 1% variation in cost between the set optimized in this manuscript and the set optimized at an SNR of 15 to 25. Because the TE-correction primarily affects simulated SNR, it is expected that the *b* value set is still relatively optimal, even if the TE used and SNR achieved are slightly different across systems. Therefore, the optimal *b* value sets that we propose could be of interest as a first step toward standardization.

Compared to the reference *b* value set of 20 *b* values, the optimal *b* value set showed superior SD in *f*, *D*, and *K*, and slightly inferior SD in *D**. This indicates that precision is improved because the underlying physiological variation was assumed to remain constant within the scan session. Interestingly, based on the relative CRLB, a higher variation of a factor 2.2 for *K* was expected in the reference set compared to the optimized set and a slightly decreased variation in *f*, *D*, and *D** (factor 0.5, 0.59, and 0.56, respectively) in the reference set compared to the optimized set. Thus, the actual optimized set performed better than expected for *D* and *f*, showing there are more factors at play than the image noise described by the CRLB. Repeatability could not be assessed

because a repeat of the reference scan was not acquired due to scan time limitations.

The repeatability of all parameters seems worse for the set of 30 *b* values than for the set of 20 *b* values. This seems counterintuitive but is likely due to longer scan time, resulting in an opportunity for motion and other effects corrupting the data that could eventually lead to worsening the parameter estimation. The registration method that we applied was not sufficient to account for all the time-related effects. Also, no clear impairment in repeatability metrics was observed for any of the parameters when moving from a set of 20 to a set of 15 *b* values. A tradeoff between scan time and parameter repeatability was only observed when imaging with less than 15 *b* values.

Repeatability is expected to improve in patients due to higher SNR often encountered in tumors than in healthy tissue for diffusion imaging. This is because healthy tissue has a lower intravoxel water content than nonnecrotic tumor tissue. Therefore, even if the measurement variation in *f*, *D**, and *K* are high, they might become acceptable in patients. Especially *K* is expected to improve because it mainly relies on high *b* values where the increase in SNR has the highest impact.

No clear advantage is shown for applying inter- and intravolume registration in this work. This might be because volunteers generally lie very still, which would not necessarily be the case in patients. It might also be that once the estimation is based on a multi-*b* value fitting, a sufficient number of *b* values make the estimation robust enough regarding motion. However, a side effect of intravolume registration is smoothing of the *b* value images. This could lead to smoothing of underlying physiological

differences and might be problematic when assessing intra-tumor heterogeneity.

In this study, we have shown that swallowing artifact image rejection improves the accuracy of parameter estimation. Previously, Chevallier et al[30] already showed that rejection of motion corrupted volumes is beneficial for estimation of IVIM parameters. In this paper, the clearest effect is seen in the perfusion fraction, especially in set sizes 20, 15, 10, and 5. This could be because that the relative signal drop is larger in low b values ($b < 200$ s/mm^2) than in higher b values. Furthermore, our results suggest that there is a threshold for the amount of artifacts of around 10% that need to occur before swallowing artifact image rejection is worthwhile. In set size 6 and 30, no strong swallowing artifact image rejection effect was observed, whereas in a set size of 5, 10, 15, and 20, more than 10% of data was rejected and a clear effect is seen in the perfusion fraction.

Because swallowing artifacts can severely hamper parameter estimation, swallowing artifact rejection is beneficial. An automated strategy for swallowing artifact rejection, such as proposed by Gurney-Champion et al,[31] could be a relatively simple way to implement this procedure in clinical practice. As a topic for further research, the effect of swallowing artifact correction in the head and neck region could be investigated, for example, using a combined principal component analysis and temporal maximum intensity projection approach PCATMIP.[32]

The main limitation of this work is that it was performed on healthy volunteers for whom the SNR at high b values was relatively low compared to tumor tissue. Another limitation is that a correction for multiple comparisons was not applied. For these 2 reasons, the results should be independently validated in a larger cohort with head and neck cancer patients.

5 | CONCLUSION

The effect of b value optimization, protocol efficiency, registration, and swallowing artifact image rejection on parameter precision of hybrid IVIM-DKI was assessed. Optimization of b values is recommended because it improved the precision of 3 (D, f, K) out of 4 parameters of interest compared to the reference set. The b value set of 15 images ($b = 1 \times 0, 1 \times 10, 2 \times 80, 1 \times 130, 1 \times 570, 2 \times 770, 2 \times 780, 1 \times 790,$ and 4×1500) yielded the optimal tradeoff between scan time and parameter precision, with a repeatability comparable to the set of 30 in half the scan time. No clear advantage of image registration was demonstrated. However, swallowing artifact image rejection was beneficial when more than roughly 10% of the images contained artifacts.

ORCID
Nienke D. Sijtsema https://orcid.org/0000-0001-7928-1578
Dirk H. J. Poot https://orcid.org/0000-0003-0656-2963

REFERENCES
1. 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;20:1703-1714.
2. Noij DP, Martens RM, Marcus JT, et al. Intravoxel incoherent motion magnetic resonance imaging in head and neck cancer: a systematic review of the diagnostic and prognostic value. *Oral Oncol*. 2017;68:81-91.
3. Lu Y, Jansen JFA, Mazaheri Y, Stambuk HE, Koutcher JA, Shukla-Dave A. Extension of the intravoxel incoherent motion model to non-gaussian diffusion in head and neck cancer. *J Magn Reson Imaging*. 2012;36:1088-1096.
4. Fujima N, Sakashita T, Homma A, Yoshida D, Kudo K, Shirato H. Utility of a hybrid IVIM-DKI model to predict the development of distant metastasis in head and neck squamous cell carcinoma patients. *Magn Reson Med Sci*. 2018;17:21-27.
5. Rosenkrantz AB, Padhani AR, Chenevert TL, et al. Body diffusion kurtosis imaging: Basic principles, applications, and considerations for clinical practice. *J Magn Reson Imaging*. 2015;42:1190-1202.
6. Leibfarth S, Winter RM, Lyng H, Zips D, Thorwarth D. Potentials and challenges of diffusion-weighted magnetic resonance imaging in radiotherapy. *Clin Transl Radiat Oncol*. 2018;13:29-37.
7. Bito Y, Hirata S, Yamamoto E. Optimum gradient factors for apparent diffusion coefficient measurements. In Proceedings of the *SMR/ESMRMB Joint Meeting*. Nice, France; 1995. p. 913.
8. Xing D, Papadakis NG, Huang CLH, Lee VM, Carpenter TA, Hall LD. Optimised diffusion-weighting for measurement of apparent diffusion coefficient (ADC) in human brain. *Magn Reson Imaging*. 1997;15:771-784.
9. Brihuega-Moreno O, Heese FP, Hall LD. Optimization of diffusion measurements using Cramer-Rao lower bound theory and its application to articular cartilage. *Magn Reson Med*. 2003;50:1069-1076.
10. Peña-Nogales O, Hernando D, Aja-Fernandez S, de Luis-Garcia R. Determination of optimized set of b-values for apparent diffusion coefficient mapping in liver diffusion-weighted MRI. *J Magn Reson*. 2020;310:106634.
11. Kingsley PB, Monahan WG. Selection of the optimum b factor for diffusion-weighted magnetic resonance imaging assessment of ischemic stroke. *Magn Reson Med*. 2004;51:996-1001.

12. Saritas EU, Lee JH, Nishimura DG. SNR dependence of optimal parameters for apparent diffusion coefficient measurements. *IEEE Trans Med Imaging*. 2011;30:424-437.

13. Merisaari H, Jambor I. Optimization of b-value distribution for four mathematical models of prostate cancer diffusion-weighted imaging using b values up to 2000 s/mm(2): simulation and repeatability study. *Magn Reson Med*. 2015;73:1954-1969.

14. Leporq B, Saint-Jalmes H, Rabrait C, et al. Optimization of intra-voxel incoherent motion imaging at 3.0 Tesla for fast liver examination. *J Magn Reson Imaging*. 2015;41:1209-1217.

15. Zhang JL, Sigmund EE, Rusinek H. et al. Optimization of b-value sampling for diffusion-weighted imaging of the kidney. *Magn Reson Med*. 2012;67:89-97.

16. Lemke A, Stieltjes B, Schad LR, Laun FB. Toward an optimal distribution of b values for intravoxel incoherent motion imaging. *Magn Reson Imaging*. 2011;29:766-776.

17. Jalnefjord O, Montelius M, Starck G, Ljungberg M. Optimization of b-value schemes for estimation of the diffusion coefficient and the perfusion fraction with segmented intravoxel incoherent motion model fitting. *Magn Reson Med*. 2019;82:1541-1552.

18. Wang X, Hickernell FJ. Randomized Halton sequences. *Math Comput Model*. 2000;32:887-899.

19. Lu Y, Jansen JFA, Stambuk HE, et al. Comparing primary tumors and metastatic nodes in head and neck cancer using intravoxel incoherent motion imaging: a preliminary experience. *J Comput Assist Tomogr*. 2013;37:346-352.

20. Paudyal R, Oh JH, Riaz N, et al. Intravoxel incoherent motion diffusion-weighted MRI during chemoradiation therapy to characterize and monitor treatment response in human papillomavirus head and neck squamous cell carcinoma. *J Magn Reson Imaging*. 2017;45:1013-1023.

21. Marzi S, Piludu F, Vidiri A. Assessment of diffusion parameters by intravoxel incoherent motion MRI in head and neck squamous cell carcinoma. *NMR Biomed*. 2013;26:1806-1814.

22. Alexander DC. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *Magn Reson Med*. 2008;60:439-448.

23. Andersson JLR, Skare S, Ashburner J. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*. 2003;20:870-888.

24. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004;23:S208-S219.

25. Guyader JM, Bernardin L, Douglas NHM, Poot DHJ, Niessen WJ, Klein S. Influence of image registration on apparent diffusion coefficient images computed from free-breathing diffusion MR images of the abdomen. *J Magn Reson Imaging*. 2015;42:315-330.

26. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29:196-205.

27. Bland JM, Altman DG. Measurement error. *BMJ*. 1996;313:744.

28. Poot DHJ, den Dekker AJ, Achten E, Verhoye M, Sijbers J. Optimal experimental design for diffusion kurtosis imaging. *IEEE Trans Med Imaging*. 2010;29:819-829.

29. Kolff-Gart AS, Pouwels PJW, Noij DP, et al. Diffusion-weighted imaging of the head and neck in healthy subjects: reproducibility of ADC values in different MRI systems and repeat sessions. *Am J Neuroradiol*. 2015;36:384-390.

30. Chevallier O, Zhou N, He J, Loffroy R, Wang YXJ. Removal of evidential motion-contaminated and poorly fitted image data improves IVIM diffusion MRI parameter scan-rescan reproducibility. *Acta Radiologica*. 2018;59:1157-1167.

31. Gurney-Champion OJ, Froeling M, Klaassen R, et al. Minimizing the acquisition time for intravoxel incoherent motion magnetic resonance imaging acquisitions in the liver and pancreas. *Invest Radiol*. 2016;51:211-220.

32. Pai VM, Rapacchi S, Kellman P, Croisille P, Wen H. PCATMIP: enhancing signal intensity in diffusion-weighted magnetic resonance imaging. *Magn Reson Med*. 2011;65:1611-1619.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

FIGURE S1 Overview of the workflow for obtaining the difference in mean and coefficient of repeatability (CR). Each acquisition (denoted A and B), contained multiple b-values. After distortion correction and registration, a fit can be made for each voxel, yielding estimates of the 4 IVIM-DKI parameters. Subsequently, the mean in the tonsils was calculated over both tonsils for each parameter for acquisitions A and B and the mean of acquisition A was subtracted from the mean of acquisition B to yield d_j. These steps were repeated for each volunteer. Finally, the average difference in mean and CR were calculated

FIGURE S2 (A) Axial T2-weighted slice with (B) corresponding $b = 130$ s/mm^2 axial slice with the tonsils outlined in red and (C) identical slice to B, but affected by a swallowing artifact. In (D) the same slice is shown but for $b = 1500$ s/mm^2. In the bottom row parametric maps for the tonsils are shown as color overlays on the corresponding $b = 0$ s/mm^2 slice with in (E) the perfusion fraction f, in (F) pseudo-diffusion coefficient $D*$ in mm^2/s, in (G) diffusion coefficient D in mm^2/s and in (H) kurtosis K

FIGURE S3 Boxplots showing the distribution of the mean and standard deviation (SD) in the tonsils over the volunteers for each of the four parameters of interest; the perfusion fraction f, the pseudo-diffusion coefficient $D*$, the diffusion coefficient D and the kurtosis K. The optimized set of 20 b-values is marked "O" and the reference set "R". *P*-values are indicated in cas$E$ $P \leq .05$ according to the Wilcoxon signed rank test

FIGURE S4 (A) Relative Coefficient of repeatability (CR) and (B) relative difference in mean between two identical acquisitions shown as a percentage of the overall mean of the two acquisitions for the four parameters of interest; the perfusion fraction f, the pseudo-diffusion coefficient $D*$, the diffusion coefficient D and the kurtosis K. Significant difference in mean based on Wilcoxon signed rank test was found in f for set size 10 and $D*$ for set size 20. These points are denoted by an asterisk marker. For and explanation of the metrics, see Figure 1

FIGURE S5 (A) Relative Coefficient of repeatability (CR) and (B) relative difference in mean between two identical acquisitions shown as a percentage of the overall mean of the two acquisitions for the four parameters of interest; the perfusion fraction f, the pseudo-diffusion coefficient $D*$, the diffusion coefficient D and the kurtosis K. Full registration comprises intra- and inter-volume registration. Significant difference in mean based on Wilcoxon signed rank test was found in f for no registration and inter-volume only registration. These points are denoted by an asterisk marker. For an explanation of the metrics, see Figure 1

FIGURE S6 Boxplot of the mean (A) and standard deviation (B) of f (perfusion fraction), over 16 volunteers for each set size of 5-30 b-values. The acquisition including swallowing artifacts (darkest grey) performs worse than the same acquisition after swallowing artifact correction (middle grey) when compared to the baseline acquisition (lightest grey) which did not contain any swallowing artifacts