Apparent exchange rate (AXR) mapping in diffusion MRI: An in vivo test-retest study and analysis of statistical power.

Björn Lampinen[1], Filip Szczepankiewicz[1], Danielle van Westen[2], Pia Sundgren[2], Freddy Ståhlberg[1,2], Jimmy Lätt[3], and Markus Nilsson[4]

[1]*Dpt. of Medical Radiation Physics, Lund University, Lund, Sweden,* [2]*Dpt. of Diagnostic Radiology, Lund University, Lund, Sweden,* [3]*Center for Medical Imaging and Physiology, Lund University Hospital, Lund, Sweden,* [4]*Lund University Bioimaging Center, Lund University, Lund, Sweden*

Introduction: Diffusion weighted imaging (DWI) can be sensitized to water exchange across cellular membranes using double pulsed field gradient (d-PFG) sequences[1]. Filter-exchange imaging (FEXI) is a d-PFG method designed to gain imaging speed by restricting the analysis to the apparent exchange rate (AXR)[2]. The present study aims to prepare FEXI for future clinical research studies investigating differences in mean AXR across populations. In order to establish the number of subjects required in such studies, this work presents *in vivo* test-retest data, acquired with an optimized protocol[3], and assesses the mean AXR (μ), coefficient of variation (CV), and the relative measurement contribution to variance (RV_m) in various brain structures. We show that with group sizes below 10 subjects per group, mean differences of 40-80% in AXR are detectable.

Theory: FEXI is based on a two-component exchange model. A d-PFG sequence is used to filter out the fast water component and reduce the apparent diffusion coefficient (ADC) after the first PGSE block. Exchange takes place during the mixing time, and the apparent exchange rate is probed by measuring the partially restored ADC at the time of the second PGSE block[2,4]. For group comparisons of a parameter according to pathology, the effect of interest is often a difference in mean between two groups. In the case of AXR, this would involve a patient group and a control group, with the effect size of $\mu_1 - \mu_2 = \Delta\mu$. The statistical power of the comparison increases with larger group sizes and smaller parameter variance. For AXR, the variance can be regarded as arising from a two-level random effects model, where an individual AXR value (Y) is expressed in terms of the population mean (μ) offset by an inter-subject error (ε_i) and a measurement error (ε_m), $Y = \mu + \varepsilon_i + \varepsilon_m$. Assuming those errors are unbiased and independent, the total variance (V) is then the sum of the inter-subject variance (V_i) and the measurement variance (V_e), according to $V = V_i + V_e$. In this study, we considered the total variance in terms of the coefficient of variation (CV), defined as the standard deviation ($\sigma = V^{1/2}$) normalized by the mean, $CV = \sigma/\mu$. We also analyzed the relative measurement contribution to variance, $RV_m = V_e/V$. This last measure relates to how much the statistical power increases when the measurement is improved by taking actions such as upgrading the gradient performance, adjusting the protocol, prolonging the scan time et cetera[5].

Method: Eighteen subjects were scanned using a Philips Achieva 3T system. Two subsequent FEXI acquisitions were performed, each with a scan time of 13 minutes, together with a whole-brain DTI-scan. FEXI data was acquired over seven slices at a spatial resolution of 3×3×5 mm^3 and TE/TE$_f$/TR = 68/38/2500 ms. An optimized protocol was used, applying diffusion encoding in six directions, where b_f = 830 s/mm^2 for the filter block, and $\boldsymbol{b}$ = 40 and 1300 s/mm^2, repeated three and six times, respectively, for the detection block[3]. For each combination of b-values, acquisitions were performed with two mixing times: t_m = 16 and 442 ms, each repeated twice. During one of the short mixing times, the filter block was inactive. In the obtained parameter maps, five bilateral regions of interest (ROI) were drawn, based on directionally color encoded FA-maps: the anterior corona radiata (ACR), the anterior limb of the internal capsule (ALIC), the corpus callosum (CC), the cerebrospinal tract (CST) and the thalamus (TH) (Fig. 1 c and d). Each ROI was analyzed for μ, CV and RV_m. Also, power calculations were performed to determine the group size required to achieve a statistical power of 0.8, at a level of significance of 0.05, assuming equal sized t-distributed samples from two populations, normally distributed with different means but equal variance. The analysis postulated various effect sizes as defined by a modified Cohen's d, taken as $\Delta\mu/V_i^{1/2}$.

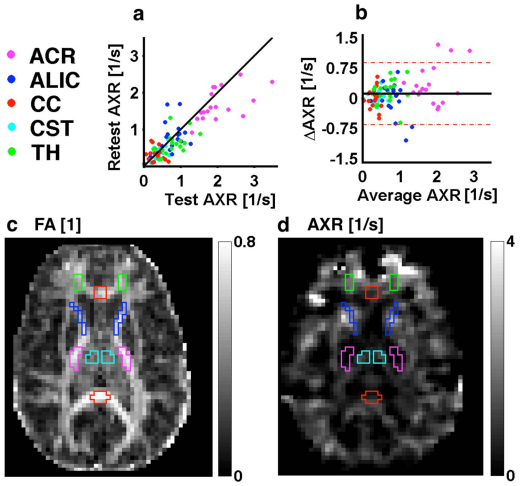

Fig. 1
FEXI test-retest results. The scatter plot (a) shows inter-subject variance parallel to the line and measurement variance perpendicular to it. In the Bland-Altman plot (b) there is small test-retest disagreement, which is significant at approximately 19% in the ACR. Below, the ROI used in the study are shown for one of the subjects in an FA map (c) and an AXR map (d).

Table 1.
The table specifies the AXR mean (μ in s^{-1}) with the standard error of the mean (SE), the coefficient of variation (CV in %), the relative measurement contribution to variance (RV_m in %), and the mean number of voxels for each ROI. In the ALIC and CC, the variance is mostly due to the measurement, but in the ACR, CST and TH, inter-subject differences account for approximately half of the observed variability.

	AXR			
Region	$\mu \pm SE$	CV	RV_m	mean #$_{voxels}$
ACR	1.90 ± 0.10	27	46	29
ALIC	0.81 ± 0.06	41	81	90
CC	0.31 ± 0.03	58	108	87
CST	0.37 ± 0.04	45	31	42
TH	0.69 ± 0.04	46	48	64

Results: Results of the test-retest measurements are presented in Table 1, including CV for AXR in the range of 27 to 58%. Concerning RV_m, it is high for ALIC and CC, however, judging by the RV_m values for the ACR, CST and TH, approximately half of the variance in these regions is due to inter-subject differences. The inter-subject variance component is also illustrated by the distributions of same-colored points parallel to the line in the scatter plot and Bland-Altman plot in Fig. 1 (a and b). In the Bland-Altman plot, a small test-retest disagreement can be observed, significant only in the ACR at approximately 19%. Power calculations show that with a Cohen's d of 1, 2 and 3 in the CST (i.e. $\Delta\mu$ of approximately 0.14, 0.27 and 0.41 s^{-1}, or 38, 76 and 114% from the mean) required group sizes would be 24, 7 and 4 subjects, respectively.

Discussion and conclusions: AXR values seen in this study are consistent with those in earlier reports, being high in the frontal white matter (ARC), around unity in the internal capsule (ALIC) and lower in the deep gray matter (TH)[4]. The CV of AXR is approximately ten times larger than that of the ADC values estimated from the same data. Factors contributing to high AXR variance likely include the low SNR of d-PFG sequences and noise sensitivity inherent to the model. Also, in the ACR, CST and TH approximately half of the variance was due to inter-subject differences. This increases the reliability of follow-up tests on the same individual, but also put a natural limit on the potential gain from stronger gradients or longer scan time. Lastly, the use of a simple two-level random effects model may limit the quantification of variance contribution. We expect that more complete models can account for the overestimation of RV_m in CC, although the main findings of this work are still valid. In conclusion, studying groups below 10 subjects, effect sizes approximately twice the inter-subject standard deviation could be detected, corresponding to an alteration of 40-80% in AXR, depending on the region.

References:

[1]Lätt *et al.* NMR Biomed 2009;22:619–628 [2]Lasiĉ *et al.*MRM 2011;66:356-365 [3]Lampinen *et al.* ISMRM 2013; 0497 [4]Nilsson, *et al.* MRM 2012;doi:10.1002/mrm.24395 [5]Szczepankiewicz *et al.* NeuroImage 2013;76:145-154