Evaluating Quantitative Proton-Density-Mapping Methods

Aviv Mezer,[1]* Ariel Rokem,[2] Shai Berman,[1] Trevor Hastie,[3] and Brian A. Wandell[4,5]

[1]*The Hebrew University of Jerusalem, Edmond and Lily Safra Center for Brain Sciences, Jerusalem, Israel*
[2]*The University of Washington, eScience Institute, Seattle, WA, USA*
[3]*Stanford University, Department of Psychology, Stanford, CA, USA*
[4]*Stanford University, Department of Psychology, Stanford, CA, USA*
[5]*Stanford University, Center for Cognitive and Neurobiological Imaging, Stanford, CA, USA*

Abstract: Quantitative magnetic resonance imaging (qMRI) aims to quantify tissue parameters by eliminating instrumental bias. We describe qMRI theory, simulations, and software designed to estimate proton density (PD), the apparent local concentration of water protons in the living human brain. First, we show that, in the absence of noise, multichannel coil data contain enough information to separate PD and coil sensitivity, a limiting instrumental bias. Second, we show that, in the presence of noise, regularization by a constraint on the relationship between T1 and PD produces accurate coil sensitivity and PD maps. The ability to measure PD quantitatively has applications in the analysis of in-vivo human brain tissue and enables multisite comparisons between individuals and across instruments. *Hum Brain Mapp 00:000–000, 2016.* © 2016 Wiley Periodicals, Inc.

Key words: quantitative magnetic resonance imaging; proton density; T1; coil sensitivity; parallel imaging

INTRODUCTION

Quantitative magnetic resonance imaging (qMRI) methods seek to measure biophysical properties of the substrate. QMRI measurements are of great value because they can be compared in a single subject across different instruments at multiple time points, as well as between different subjects.

Proton density (PD) is the most basic MRI measure, representing the apparent concentration of water protons (mobile hydrogen atoms) in each voxel. Water concentration differs between brain tissue types, and the general ability to infer quantitative tissue properties from qMRI measurements requires knowledge of the ratio between macromolecules and water within each voxel [Tofts, 2003].

Several PD-estimation techniques have been described [Abbas et al., 2014, 2015; Mezer et al., 2013; Neeb et al., 2006; Noterdaeme et al., 2009; Volz et al., 2012a,b; Whittall et al., 1997]. These techniques quantify PD by accounting for coil-sensitivity maps (receive inhomogeneity) in three different ways. First, some techniques combine data from multichannel coils into a single channel while others keep data from the multiple channels separate during the estimation. Second, the techniques use different regularizing assumptions to overcome the ill-posed nature of the

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Aviv Mezer, Edmond and Lily Safra Center for Brain Sciences (ELSC), The Hebrew University of Jerusalem, Jerusalem, 91904, Israel. E-mail: Aviv.Mezer@ELSC.huji.ac.il

Received for publication 24 September 2015; Revised 30 April 2016; Accepted 10 May 2016.

DOI: 10.1002/hbm.23264

Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com).

problem. Third, the techniques differ as to whether the method relies on a single global brain analysis, or a set of local calculations that are integrated in a final step.

Here, we describe, implement, and analyze these PD-estimation techniques for the purpose of comparing their efficacy and accuracy. In Theory, we explain the general principles and why the estimation is ill-posed, requiring regularization. We then explain several different regularization approaches. Finally, we contrast local versus global estimation methods. In Methods, we describe the simulation, acquisition and fitting pipelines. In Results, we compare the theory and measurements. We support [Volz et al., 2012b]'s finding that it is possible to achieve high PD-estimation accuracy using T1 regularization, and we further show that the best results are obtained by using local T1 regularization based on many small, overlapping volumes.

THEORY

We introduce the fundamental MRI equations for PD estimation from a spoiled-gradient-echo (spoiled-GE) MRI sequence. We explain why separating PD from coil sensitivity is ill-posed. We then systematically describe solutions to this problem.

MRI Signal Equations for PD Estimation

The measured signal from a voxel in any brain spoiled-GE image depends on $\mathrm{M0}(x,y,z)$, which is the Hadamard product ($\circ$) of the receive-coil sensitivity ($G(x,y,z)$) and the proton density ($\mathrm{PD}(x,y,z)$) of the brain tissue [Eq. (1)].

$$\mathrm{M0}(x,y,z)=G(x,y,z)\circ \mathrm{PD}(x,y,z)= G\circ \mathrm{PD} \tag{1}$$

The signal equation for the spoiled-GE MRI sequence, showing the combination of brain and instrumental parameters, is:

$$S(\alpha) = \left(G \circ \mathrm{PD}\right)\ e^{-\left(\frac{\mathrm{TE}}{\mathrm{T2^*}}\right)}\sin(\alpha)\ \left(\frac{1-e^{-\left(\frac{\mathrm{TR}}{\mathrm{T1}}\right)}}{1-\cos(\alpha)e^{-\left(\frac{\mathrm{TR}}{\mathrm{T1}}\right)}}\right) \tag{2}$$

Quantitative MRI aims to separate the brain substrate from the instrumental factors (coil sensitivity, also called coil gain). Equation (2) shows that T1 can be estimated with multiple measurements using at least two different flip angles (α) [Deoni et al., 2003; Fram et al., 1987].

Equation (2) also shows that the PD estimate is linked with the spatial variations in coil sensitivity (G), and that no simple imaging manipulation can separate PD from G.

To see this, note that we can insert an arbitrary matrix, A and the $\widehat{A}$ (the Hadamard inverse of A), into the Eq. (1) without changing the measured signal (i.e., $[\mathrm{M0}= G\circ \mathrm{PD} = G\circ \mathrm{A}\circ \widehat{A}\circ \mathrm{PD}]$) . Hence, $(G\circ \mathrm{A})$ and $(\widehat{A}\circ \mathrm{PD})$ are also solutions, and the equation can only be solved for PD using additional constraints.

Furthermore, after separating G and PD in space, one needs to find a single scalar that calibrates the PD values to establish physical units. Here, we set the scale using the common assumption that in brain ventricles $\mathrm{PD}=1$ ("Methods," Scaling PD to Water Content section).

Constraining the Ill-Posed PD-Estimation Problem

We analyze five approaches that contribute to solving the ill-posed problem: (a) the coil-sensitivity functions are smooth over the volume, (b) the information in the biophysical relationship between PD and T1, (c) the additional information available by analyzing the coil channels separately, (d) the value of further regularization principles, and (e) building the solution from many local estimates rather than a single global estimate.

Smoothness constraints

A general approach to separating PD and G is to assume that the coil-sensitivity functions are smooth over space [Abbas et al., 2015; Blaimer et al., 2004; Mezer et al., 2013; Noterdaeme et al., 2009; Volz et al., 2012a,b]. We can implement this constraint by approximating the coil sensitivity as low-order 3D-polynomial functions specified by parameters $< p >$.

A K^{th}-order 3D polynomial can be express by Eq. (3), where $i+j+k \le K$ and i, j, and i are non-negative.

$$p_0+\sum_{i,j,k} p_{ijk}x^i y^j z^k \tag{3}$$

For example, a first-order (linear $K=1$) polynomial has four parameters [Eq. (3a)]

$$G_{\mathrm{p}}(x,y,z)=p_0+p_1x+p_2y+p_3z \tag{3a}$$

A second-order polynomial ($K=2$) has 10 parameters [Eq. (3b)]

$$\begin{aligned} G_{\mathrm{p}}(x,y,z)=p_0+p_1x+p_2y+p_3z+p_4x^2+p_5y^2+p_6\ z^2 \\ +p_7xy+p_8yz+p_9zx \end{aligned} \tag{3b}$$

The smoothness constraint is not sufficient to solve for the PD. Suppose that the number of polynomial coefficients is

Abbreviations

MAPE	The median of the absolute percent errors
NLS	Nonlinear least squares
PD	Proton density
qMRI	Quantitative magnetic resonance imaging
SEIR	Spin echo inversion recovery
s.d.	Standard deviation
Spoiled-GE	Spoiled gradient echo

N_p, and we measure the N_v voxels with a single channel. There are N_v unknown PD values and N_p unknown coil sensitivity parameters. There are more unknowns ($N_v + N_p$) than measurements (N_v). Hence, calculating a unique PD requires additional assumptions.

One option is to assume that all of the smooth variation in the M0 image is explained by the coil-sensitivity function G_p [Noterdaeme et al., 2009; Volz et al., 2012a]. But this approach is limited because true PD also contains smooth variation. Below, we describe alternative sources of information.

T1 constraints

[Volz et al., 2012b] found that both theory and observation confirm that, in any given tissue type, there is a relationship between T1 and PD, and showed that it can be used to separate PD from G. A large part of the T1 variation is explained by the concentration of non-water tissue [Kim et al., 1994; Kucharczyk et al., 1994; Stüber et al., 2014]. Across the conditions used for human brain imaging, the relationship between 1/T1 and 1/PD is linear [Eq. (4)], [Abbas et al., 2015; Fatouros and Marmarou, 1999; Gelman et al., 2001; Mezer et al., 2013; Tofts, 2003].

$$\frac{1}{\mathrm{PD}_{\mathrm{pred}}} = \frac{\gamma}{T1} + \delta; \;\; \mathrm{PD}_{\mathrm{pred}} = \frac{T1}{\gamma + \delta T1} \tag{4}$$

Multichannel information

[Mezer et al., 2013] use multichannel data to separate PD from G. Specifically, data from multiple channels provide an over-determined system of equations. Suppose that the number of polynomial coefficients is N_p, the number of channels is N_c, and the number of voxels we are measuring is N_v. There are N_v unknown PD values, and $N_c \times N_p$ unknown channel-sensitivity parameters, and $N_c \times N_v$ measurements. There will be more measurements than unknowns when $N_c N_v > N_v + N_c N_p$.

If G is a polynomial of relatively low order, the number of measurements exceeds the number of unknown parameters. Figure 1 illustrates the relationship between the substrate PD (right image), the three-channel sensitivity functions (middle column), and the M0 images from three channels (left column).

Using multichannel data, a solution for PD and G can be found by minimizing an objective function that includes the difference between the observed and the predicted M0. The minimization can be calculated in two different ways. We can use an alternating least-squares (ALS) method [Eq. (5)] (λ=0) that iterates between the two linear equations, fixing either ($\hat{p}$ or $\widehat{\mathrm{PD}}$), until convergence.

$$\hat{p}, \widehat{\mathrm{PD}} = \mathrm{argmin}_{p,\mathrm{PD}} \left\{ ||M0_{0\mathrm{bs}} - M0_{\mathrm{pred}}||_2^2 + \lambda \mathrm{Err}_{\mathrm{reg}} \right\} \tag{5}$$

Alternatively, we can use a nonlinear search for $\hat{p}$ the polynomial coefficients, solving for $\widehat{\mathrm{PD}}$ as a nested least squares that is part of the minimization function [Eq. (6)] (λ=0).

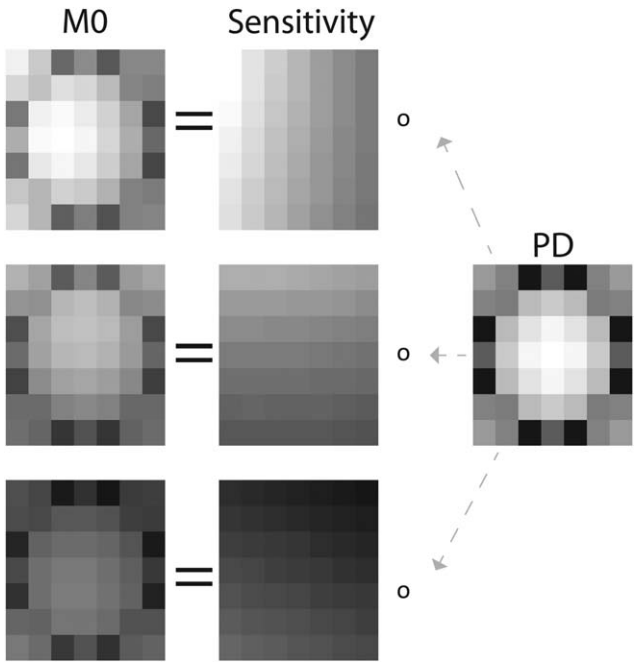

Figure 1.
The problem as a picture. The images show the relationship between proton density (PD), coil sensitivity, and the MRI M0 data [Eq. (1)]. The PD values are multiplied, point-by-point (Hadamard product), by the coil sensitivities, to produce the M0 images. In this simulation, the coil sensitivities are second-order polynomials.

$$\hat{p} = \mathrm{argmin}_p \left\{ ||M0_{0bs} - M0_{\mathrm{pred}}||_2^2 + \lambda \mathrm{Err}_{\mathrm{reg}} \right\} \tag{6}$$
$$\widehat{\mathrm{PD}} = \overline{M0_{0\mathrm{bs}}{}^{\circ} \left(\frac{1}{G_p} \right)}; \;\; M0_{\mathrm{pred}} = \widehat{\mathrm{PD}}^{\circ} \; G_p$$

In the presence of measurement noise, these solutions are subject to overfitting. To limit this possibility, we use a regularization term $\lambda \mathrm{Err}_{\mathrm{reg}}$ and cross-validation to quantify the fit's accuracy. The regularization term expresses additional constraints, permitting a slightly worse fit to the data in exchange for satisfying other important constraints. The regularization parameter λ expresses the tradeoff between fitting the data and these additional constraints.

Regularization for multichannel approach

We evaluated three regularization techniques:

Ridge (Tikhonov) regression. Ridge regression (also called Tikhonov regression) is widely used in parallel imaging [Hoge et al., 2005; Liang et al., 2002; Lin et al., 2004]. This regularizer selects a vector of coil-sensitivity polynomial coefficients (p) with a small vector length

($\mathrm{Err}_{\mathrm{reg}} = ||p||_2^2$). When an orthonormal polynomial basis is chosen, minimizing the vector length of p is equivalent to minimizing the vector length of the coil-sensitivity coefficients. Tikhonov regularization has a closed-form solution that can be obtained with ALS [Eq. (5)].

Correlation regularization. The second regularizer constrains the correlations between M0 measurements and channel sensitivity. The PD estimation is based on the fact that each channel carries different sensitivity information. The correlation between M0 measurements from any two channels, $\mathrm{corr}(M0_i,\ M0_j)$ must be greater than the correlation between the corresponding sensitivity maps of these channels $\mathrm{corr}(G_i,\ G_j)$, [Mezer et al., 2013]. This regularization reduces overfitting by limiting the correlation between the coil sensitivities ($G_i,\ G_j$), keeping the correlation below that found in the measurements ($M0_i,\ M0_j$).

$$\mathrm{Err}_{\mathrm{reg}} = \begin{cases} 1, & if\ corr(M0_i,\ M0_j) < corr(G_i,\ G_j) \\ 0, & otherwise \end{cases}$$

With correlation regularization there is no closed-form solution, and a non-liner search can be used [Eq. (6)].

Biophysical regularization. The Tikhonov and coil-correlation regularizers arise from general mathematical ideas; the third regularizer is motivated by specific biophysical information.

Volz's group [Volz et al., 2012b] suggests that PD can be accurately estimated by accounting for a global linear relationship between T1 and PD [Eq. (4)]. A similar approach can be used to derive a regularization term for the multi-channel problem [Eq. (6)]:

$$\mathrm{Err}_{\mathrm{reg}} = ||\widehat{\mathrm{PD}} - \mathrm{PD}_{\mathrm{pred}}||_2^2 = ||\widehat{\mathrm{PD}} - \frac{T1}{\gamma + \delta T1}||_2^2$$

The regularization term ($\mathrm{Err}_{\mathrm{reg}}$) is calculated as follows. Given an estimate of the sensitivities' coefficients $\hat{p}$, we calculate G_{p}, $\widehat{\mathrm{PD}}$, and $M0_{\mathrm{pred}}$. We then use T1 and $\widehat{\mathrm{PD}}$ values to estimate the parameters γ and δ as a nested-least-squares problem that minimizes the error $\mathrm{Err}_{\mathrm{reg}}$.

The parameters ($\gamma,\ \delta$) of the linear relationship between 1/T1 and 1/PD only depend on the selected voxels' T1 and $\widehat{\mathrm{PD}}$ values. The regularization penalty is imposed only for deviations from the linear relationship, with no preference for any specific ($\gamma,\ \delta$) values; the linear parameters may vary across tissue types and individuals because each tissue type contains different molecules and ion content [Kim et al., 1994; Kucharczyk et al., 1994; Mezer et al., 2013; Stüber et al., 2014].

Global versus Local Fitting

There is a third general decision we can make about the PD-estimation method. We might try to fit the entire volume (global) at once through the minimization process. A second possibility is to use a relatively low-order polynomial over many small volumes (local) [Mezer et al., 2013], and then to smoothly join these estimates (see Combining Local PD Estimates section and Supporting Information Appendix 3). These methods differ in noise sensitivity, number of unknown parameters, computational time, and ability to capture regional differences. We compare the global and local methods in Results.

METHODS

Software

We provide a full implementation (Matlab) of the image-processing, analysis and simulation code, along with test data. Software that enables the reader to reproduce the main figures in this article is available at [https://github.com/mezera/mrSensitive]. The image-processing and simulation libraries are maintained at: https://github.com/mezera/mrQ.

Human Subjects

The Stanford University Institutional Review Board approved all procedures for medical research involving human subjects. Human measurements were performed on a healthy adult volunteer who provided informed written consent. A sample data set can be found at http://purl.stanford.edu/nn554zr6949.

Phantom

A homogeneous standard agar phantom (fBIRN) was used to evaluate the quality of the coil- sensitivity inhomogeneity correction. The fBIRN standard phantom is constructed using a 17.5-cm-diameter spherical container that contains a mixture with approximately 3.6 L H2O; 400 ml 21.8 mM NiCl2; 120 grams Agar and 20 grams NaCl (0.5%) (for more details please see: http://www.birncommunity.org/tools-catalog/function-birn-stability-phantom-qa-procedures/).

A sample phantom data set can be found at [http://www.birncommunity.org/tools-catalog/function-birn-stability-phantom-qa-procedures/].

MR-Image Acquisition

Phantom and human data ($N = 15$ ages 19–38). were obtained using a 3T GE Signa 750 MRI scanner. We obtained data using a Nova 32-channel receive-only head coil. The quantitative T1 and PD parameters were measured from spoiled-GE images acquired with different flip angles (4°, 10°, 20°, 30°), and TE = 2.4 ms. The human-brain measurements were made at $0.93 \times 0.93 \times 1\ \mathrm{mm}^3$ resolution and with a TR = 14 ms; the phantom data were acquired at 2 mm isotropic resolution and TR = 20 ms.

The RF excitation flip angles were calibrated using spin echo inversion recovery (SEIR) scans as reference [Mezer

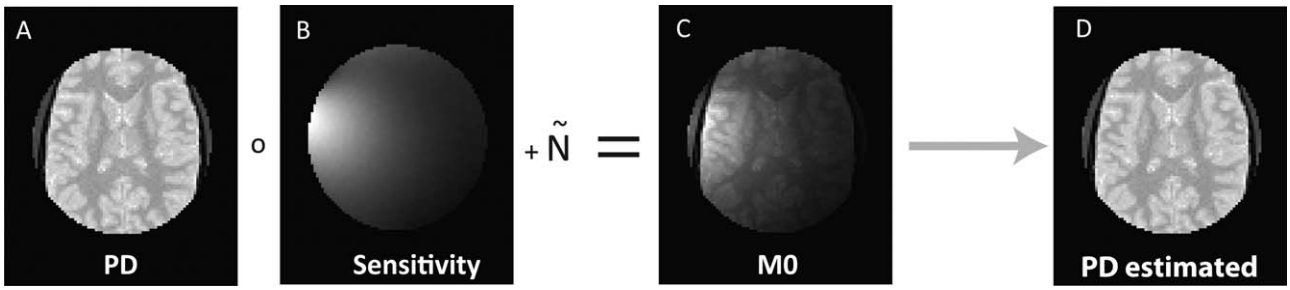

Figure 2.
A full-brain simulation comparing PD estimates. Typical human PD values (panel **A**) were multiplied by different 32-coil sensitivity functions estimated from phantom data (one shown, panel **B**). Gaussian noise (N) was added to produce simulated M0 data (panel **C**). PD was estimated from the simulated M0 data (panel **D**) using different methods (see Figs. 5 and 6).

et al., 2013]. These scans were done with TR = 3 s, slab inversion pulse and spatial spectral fat suppression. The echo time (TE) was set to minimum full; inversion times were 50, 400, 1,200, 2,400 ms. We used 2×2 mm^2 inplane resolution with a slice thickness of 4 mm. EPI readout was performed (SEIR-epi) using 2× acceleration at 3T.

The data acquisition takes about 4 min for the SEIR-epi and 8 min for spoiled-GE for $2 \times 2 \times 2$ mm^3 resolution with four flip angles, or 20 min for $1 \times 1 \times 1$ mm^3 resolution. The spoiled-GE time can be halved by using two, rather than four, flip angles (for example see [Volz et al., 2012a,b]).

To test the instrument-independence of the PD and coil-sensitivity estimates, we repeated the scanning sequence described above for one subject with a GE 8-channel receive-only head coil.

Simulated PD And Coil-Sensitivity Data

We simulated a spoiled-GE signal that is a function of coil sensitivity, PD, T1, and TR using [Eqs. (1) and (2)]. Data were simulated using flip angles of 4°, 10°, 20° and 30°, and a TR of 14 ms. There was no need to simulate the TE and T2* because of assumed short TE (∼2ms).

In some cases, independent Gaussian-noise samples were added to each simulated flip angle.

We simulated 32-channel head coil-sensitivity functions with values estimated on homogeneous phantom data (see above Estimation of the Coil sensitivity-Profile with 3D Polynomials section).

We simulated both small volumes (e.g., $15 \times 15 \times 15$ voxels) with spatial PD structures (Fig. 1) and a whole-brain volume (Fig. 2).

In the small volume (local) cases, we simulated the PD values with a variety of spatial distributions. The sensitivity functions were simulated as polynomials [Eq. (3)] with coefficients that correspond to the values estimated on a homogeneous phantom. The T1 values were simulated as a function of these PD values [Eq. (4)]. In some simulations, we also introduced specific deviations from the T1-PD formula.

The whole-brain M0 data were simulated using real sensitivity and PD values. The sensitivity functions were taken from M0 images of the homogenous phantom. To achieve a realistic whole-brain PD volume, we used fitted PD values on one subject (fitted with local T1 regularization, Local multichannel methods and regularization section). Finally, we added Gaussian noise to the simulated data. If T1 values were part of the regularization, we used the estimated T1 from the same subject.

Image Processing

RF excitation bias

The nominal RF excitation flip angles (α_N) were calibrated using the SEIR-epi data [Mezer et al., 2013]. We first measured the gold standard T1 using SEIR-epi [Barral et al., 2010]. The spoiled-GE data were registered to the SEIR-epi using ANTS, a nonlinear registration software [Avants and Gee, 2004]. Finally, we calculated the true angle (α_T), correcting for excite inhomogeneity (β); ($\alpha_T = \alpha_N \beta$) by comparing the biased spoiled-GE data and the gold standard SEIR-epi T1 values [Mezer et al., 2013]. The adjustment was made by using the spoiled-GE equation and nonlinear least squares (NLS) to estimate the true flip angle. In this analysis, we assumed that the excite inhomogeneity function was smooth over space. Therefore, we calculated the smooth inhomogeneity in the SEIR-epi space and interpolated it to the higher resolution spoiled-GE imaging space (see Supporting Information Appendix 1 for the full algorithm).

For the homogeneous phantom data, the SEIR-epi mean T1 value was applied. No registration was needed in this case. The other methods were the same.

T1 and M0 fits

The T1 and M0 maps were calculated from the measured and simulated (see below) spoiled-GE data using Eq. (2) and a fitting procedure described by [Chang et al.,

2008]. This method minimizes the difference between the data and the signal-equation predictions. For the measured MRI data, we used Eq. (2), but with the corrected flip angle (α_T) instead of the nominal flip angles (α).

Estimation of the Coil Sensitivity-Profile with 3D Polynomials

We used the agar homogeneous phantom (constant PD) to estimate the coil-sensitivity functions. For the phantom, the M0 signal measures the variation in coil sensitivity, which we modeled using a 3D orthonormal basis (G_p). The quality of the fit was estimated using different polynomial orders, and local region sizes are described in Results.

Algorithms to Separate PD And Coil Sensitivity from the M0 Image

In Theory (Constraining the ill-Posed PD-Estimation Problem and Global versus Local Fitting sections), we discussed four different approaches (local vs. global and single vs. multiple coils) to estimate the PD and the coil sensitivity from the calculated M0 data [Eq. (2)]. Here, we describe implementations of each of these approaches.

Global single-channel methods

We implemented three published global single-channel methods. [Volz et al., 2012a] described UNICORT, a method that estimates a smooth field using the SPM software to remove the coil sensitivity [Weiskopf et al., 2011]. [Noterdaeme et al., 2009] suggested COIN, which uses a different approach to estimate a smooth global field. Finally, [Volz et al., 2012b] proposed a method, PseudoT1, that uses the T1 values to estimate the PD map. We implemented UNICORT and PseudoT1, assisted by personal communication with the authors.

For single-channel analyses, we combined the M0 images from a 32-channel coil using the standard sum of squares (SOS). We also tested the methods for a single channel formed as the median of the multiple channels. In the text below, SOS was used unless specified in the text.

Global multichannel method

We tested one global multichannel approach that was developed for parallel imaging (ESPIRIT) [Uecker et al., 2014]. The software and sample data for this approach are also available at [http://www.eecs.berkeley.edu/~mlustig/Software.html].

Local multichannel methods and regularization

For the local multichannel approaches, we divided the whole-brain into multiple overlapping volumes (boxes, 1.4 cm on a side, centers separated by 7 mm; each box overlaps with 32 neighbors) [Mezer et al., 2013]. We solved for the PD in each box independently and then combined the multiple estimates into a single, large PD map (see Combining Local PD Estimates section).

We used Eq. (5) or (6) (in Theory, Multichannel information section) to estimate the PD and *G*. The M0 data were predicted from the coil-sensitivity functions and PD values. We implemented three different regularization approaches (ridge, correlation, and T1-regularization), as described in the Theory section. The optimal weight for the regularization term (λ) was set using a cross-validation approach (see Selection of the regularizer weight section in Supporting Information Appendix 2). [https://github.com/mezera/mrSensitive].

In the absence of a closed-form solution, we estimated PD and *G* using both ALS solutions [Eq. (5)] and an NLS fitting procedure [Eq. (6)] (Matlab, lsqnonlin.m [MATLAB and Optimization Toolbox Release 2014a]). The estimates were very similar, but the ALS methods took much longer. Therefore, we are reporting the NLS method.

In Supporting Information Appendix 2, we explain that the validity of each regularization approach, and its cross-validation error in small volumes (3,375 voxels), were tested via simulations. In the simulations, different (1,250 datasets) coil-sensitivity functions and PD spatial distributions were simulated. The PD spatial distributions included single spots, homogeneous regions, regions with low- and high-frequency variation, random values, and circular regions.

Local single-channel method

The Global-PseudoT1 [Volz et al., 2012b] uses all non-CSF T1 values as a single global regressor to estimate the whole-brain PD and coil sensitivity (see T1 constraints and Global single-channel methods sections). We implemented a localized version (Local-PseudoT1) using Eq. (6) with Err_{reg} of the T1 biophysical regularization (see Theory, Multichannel information section). In our implementation, the algorithm was applied separately for many local volumes, each with its own T1-PD biophysical relationship. The calculation of the local coil-sensitivity function and the combination of the local fits was similar to the other local approaches (described in Local multichannel methods and regularization and Combining Local PD Estimates sections), with one difference: only a single SOS M0 volume was used.

Combining Local PD Estimates

In the case of the global problem, we used the ventricles for scaling PD values (see Scaling PD To Water Content section); however, the multilocal problem requires that we first estimate one more unknown scalar per box. For the multichannel approaches, and the Local-PseudoT1, we used a local analysis that combined estimates across a large number of small, overlapping volumes [Mezer et al.,

2013]. Each box's PD values were estimated independently, and each estimate had one unknown scalar (see MRI Signal Equations for PD Estimation section). Since all the boxes overlap in space, the scaling factors are related and the relationship can be expressed using a set of linear equations. We solved for the relationship between the scalars by imposing consistency across the overlapping boxes (see Supporting Information Appendix 3 for the full algorithm, and a sample code in SimLinJoinBoxes.m [https://github.com/mezera/mrSensitive]).

Scaling PD to Water Content

To quantify the PD estimates, we scale the values to match the theoretically known PD and T1 value of free water in the ventricles. We identify the PD scalar in three steps. First, we identify the ventricles as those voxels in the center of the brain volume (in ac-pc space) with $4.2 < T1 < 4.7$ s. Second, we divide the sensitivity map from the data with the lowest flip-angle; these have the highest signal-to-noise ratio for the ventricles [Volz et al., 2012a,b]. Last we search for a single scale factor for those voxels that best fit the data using Eq. (2) assuming the CSF parameters are PD = 1 and T1 = 4.3 s [Hopkins et al., 1986].

Comparison between PD Fitting Algorithms

To compare the different PD fitting algorithms described in (3.8) we used the following approach

a. The whole-brain simulation M0 data (see Simulated PD And coil-Sensitivity Data section) were processed with each algorithm and the PD outputs were compared against the ground-truth PD simulation.
b. In-vivo M0 data acquired on the same subject were scanned with a 32-channels coils (see MR-Image Acquisition section) and with an 8-channels GE coil. These data were processed with different algorithms describe above (3.8) to estimate the algorithm reliability and robustness in PD estimations.

Comparing PD Mapping across Subjects

Data from 15 subjects (see MR-Image Acquisition section) were fitted for PD (see Local single-channel method section).

To calculate the mean white and gray matter PD values, each subject 's brain was segmented using FSL software software [Jenkinson et al., 2012]).

RESULTS

Assessing the whole-brain-simulation PD values (Fig. 2) using the traditional SOS M0 calculation yields poor results ($R^2 = 0.32$). Using the median channel M0 rather than the SOS is not substantially better ($R^2 = 0.52$). This reinforces the need to develop the quantitative PD-estimation algorithms analyzed below.

Coil-Sensitivity Estimation Using a Smooth Polynomial

We used the phantom M0 data to establish an appropriate polynomial degree to approximate the global and local coil sensitivities for 32-channel coils (Fig. 3A). A tenth-order polynomial fit the global coil sensitivity with less than 1% error (Fig. 3B). A third-order polynomial fit the coil sensitivity in a typical local volume (10 cm^3 = 10 ml) with an error of less than 1% (Fig. 3C).

PD Estimates from Small Volumes of Simulated M0 Data

We simulated M0 images and channel-sensitivity functions, and estimated PD values. As predicted by theory, when the sensitivity functions are polynomials (G_p), and there is no noise, multichannel data can be used to estimate the true PD values perfectly (Fig. 4A). But when M0 is contaminated by Gaussian noise, the PD, and the channel sensitivities are poorly estimated, and the returned solution explains the M0 data better than the true solution (Fig. 4B). Moreover, the residual error has spatial structure, so that the voxels with a large error are clustered together. These clusters arise because the estimated sensitivity functions are smooth.

Next, we evaluated three multichannel approaches with various types of regularization (Tikhonov [Hoge et al., 2005; Liang et al., 2002; Lin et al., 2004], correlation [Mezer et al., 2013], and local T1). All regularizations dramatically reduced the noise overfitting from median absolute percent errors (MAPE) of 30.65% with no regularization, to MAPE < 10% (Fig. A-2A). The T1-regularization approach was found to be the best (MAPE = 3.39%, Fig. 4C). This is also the only approach with no significant spatial structure in the residual error (Supporting Information Appendix 2, Fig. A-2B).

Comparing PD Estimates Using Whole-Brain Simulation with Noise

We compared global and local methods based on multichannel and single-channel data. The comparison is based on a whole-brain simulation with noise (Fig. 2).

Global single-channel approach

We evaluated three global methods that estimate the PD from single-channel M0 images: UNICORT [Volz et al., 2012a], COIN [Noterdaeme et al., 2009] and PseudoT1 [Volz et al., 2012b]. The accuracy (R^2) for the three algorithms is 0.9, 0.47, and 0.92, respectively (Fig 5B).

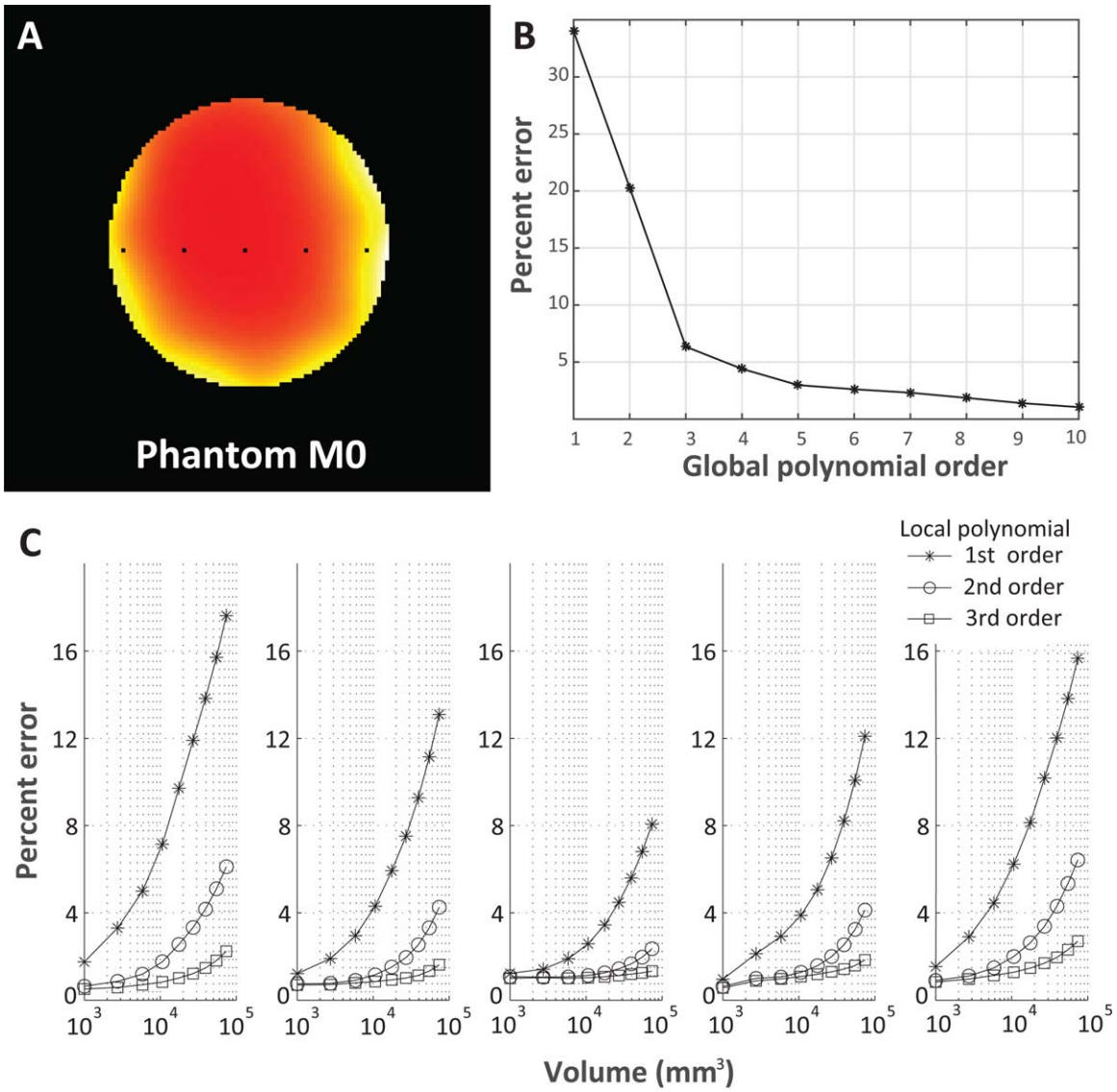

Figure 3.

Global and local polynomial fits to phantom M0 data. (**A**) The coil sensitivity was estimated from M0 data obtained from a homogenous phantom. The illustration shows 32-channel coil M0 data combined into a single channel using the standard sum-of-squares (SOS) method. (**B**) The graph shows the percent error between the whole combined M0 volume and a 3D polynomial approximation for a range of polynomial orders. A third-order polynomial has about a 7% error, and a tenth-order polynomial has a 1% error. (**C**) Local polynomial fits to phantom M0 data. The five subpanels show the percent error in the estimated multiple-channel sensitivities as a function of volume. Each subpanel shows calculations of different local volumes that are centered at each of the black points in panel A. The different symbols show the error for polynomials of different orders. In the center of the phantom, the polynomial fits do well even at larger volumes. The sensitivity changes rapidly at the edges, and the polynomial fits do well only over smaller volumes. In all cases, the error of the third-order polynomial approximation is less than ~1% for a volume of about 10^4 mm^3 (~2.15 cm isotropic or 10 ml).

Importantly, the MAPE values of each method also differ in space (Fig. 5B inserts). In general, the global approaches are very sensitive to the PD spatial structure, especially when the PD contains a smooth variation. Introducing a smooth PD variation to the global calculations in Fig. 5 considerably reduces the performance ($R^2 < 0.63$). The algorithms' accuracy also depends on how the channel data are combined, such that the combination using

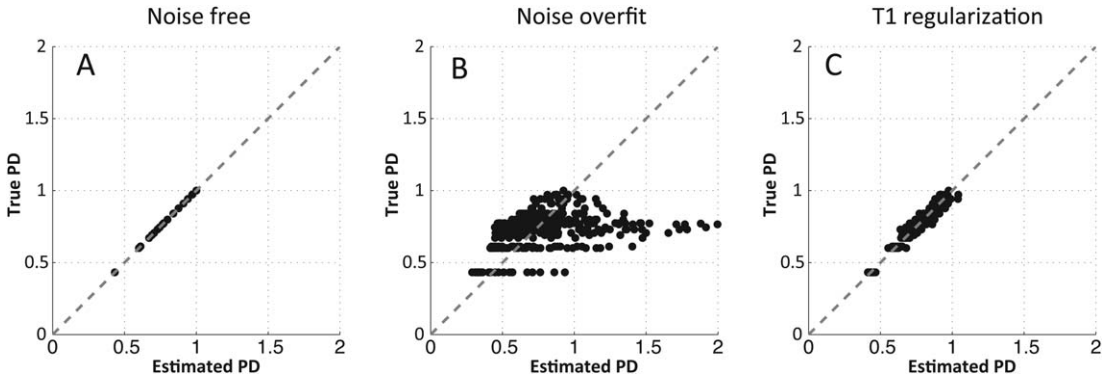

Figure 4.

Accuracy of PD estimates from multichannel data. (**A**) In the noise-free case, the M0 data from multiple coils perfectly estimates PD and polynomial coil-sensitivity functions. (**B**) In the presence of noise, the PD estimate without regularization is poor. (**C**) The PD estimate can be regularized by biophysical constraint of the T1 values. Even in the presence of substantial noise, the T1-regularized PD estimate is close to the simulated true value.

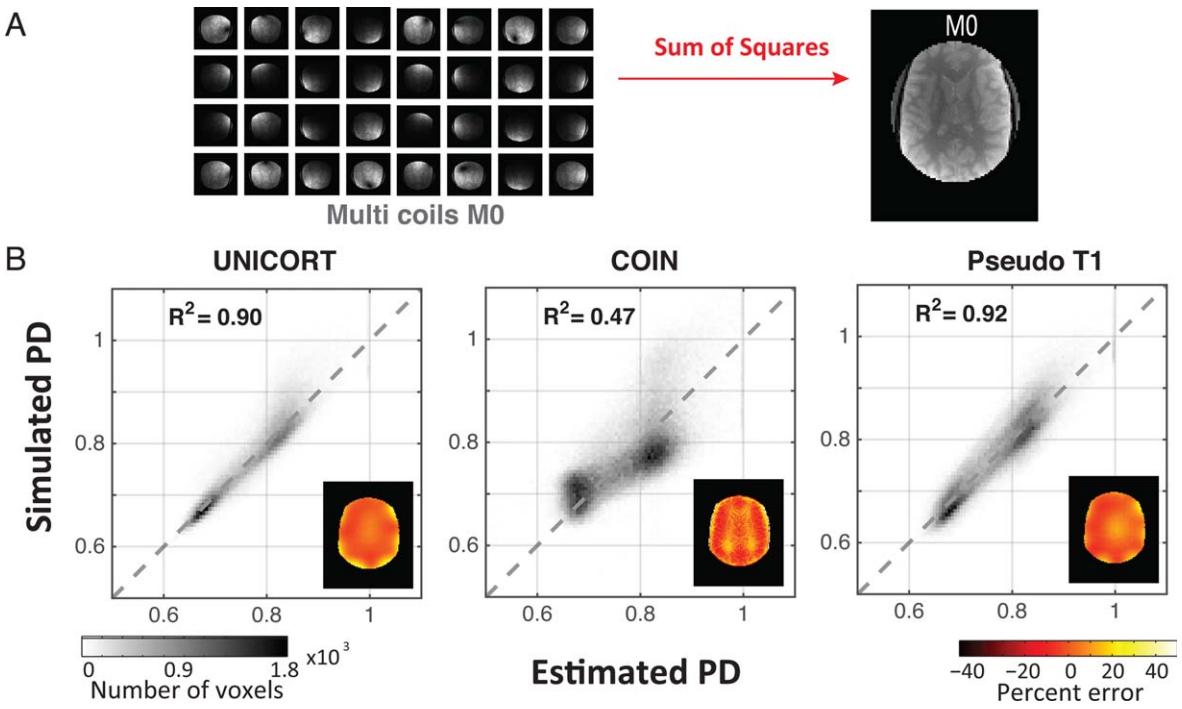

Figure 5.

Comparing full-brain PD estimates of three global approaches. (**A**) Simulated multichannel M0 volumes are combined to produce a single-channel SOS M0 volume. (**B**) The 2D histograms compare the simulated and estimated PD using three different global methods (UNICORT [Volz et al., 2012a]; COIN [Noterdaeme et al., 2009]; PseudoT1 [Volz et al., 2012b]). The image gray scale measures the number of voxels. The mean accuracy (R^2) is shown in each panel. When the data are combined using the median channel value, rather than SOS, the mean accuracy is lower (0.84, 0.47, 0.70, respectively, not shown). The inset images show the error's spatial distribution (percent).

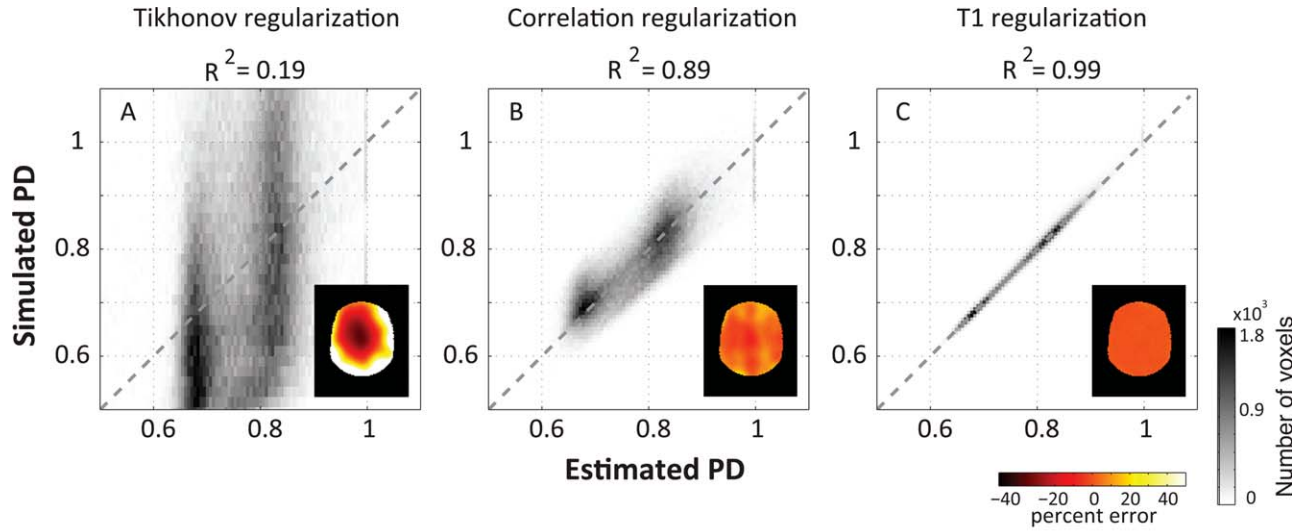

Figure 6.

Comparing full-brain PD estimates of three multichannel local approaches. The 2D histograms (panels **A–C**) compare the simulated (see Fig. 2) and estimated PD using multichannel data and the local-fitting approach for three different regularization techniques (Tikhonov, correlation and T1). The gray scale measures the number of voxels. The mean accuracy (R^2) is shown above each panel. The spatial distribution of the percent error is shown by the colored insets.

the median channel value differs significantly from the SOS combination.

Finally, the high polynomial orders suggested by [Volz et al., 2012b] for the PseudoT1 method either failed to converge or produced results with a large error. The best polynomial order was $K = 3$, which is lower than the polynomial order required to accurately fit the data ($K = 10$; Fig. 3).

Global multichannel approach

To analyze global smooth function estimated from multiple-channel-coil information, we used methods that were developed for parallel imaging (ESPIRIT) [Uecker et al., 2014]. The accuracy of this method is similar to that of the SOS value of M0 (see ESPIRIT script Example.m [https://github.com/mezera/mrSensitive]). We note that this method was not proposed for PD estimation.

Local multichannel approach

We tested three types of local multichannel regularizations: Tikhonov [Hoge et al., 2005; Liang et al., 2002; Lin et al., 2004], correlation [Mezer et al., 2013], and local T1 (Fig. 6). The Tikhonov regularization did not improve the accuracy, $R^2 = 0.19$ and the MAPE = 20% (panel A), the correlation regularization was $R^2 = 0.89$ and MAPE = 3.8% (panel B), and the local T1-regularization method recovered the PD values with high precision, $R^2 = 0.99$ and MAPE = 0.5% (panel C). The Tikhonov error and correlation error have significant spatial structure (Supporting Information Appendix 2, Fig. A-2B), and this produces low frequency errors across the brain (Fig. 6AB, insets). The T1-regularization error has substantially less spatial structure, so that accumulating data across the larger volume produces far less low-frequency spatial-coherence error (Fig. 6C inset).

Initializing the multiple-channel fitting from very different starting points had almost no effect on the fit accuracy. Furthermore, in most of the brain volume, the regularization weight selected by the cross-validation procedure was the same as the weight selected for the phantom data and phantom simulation.

The multichannel analyses are sensitive to noise. In these simulations, changing the number of channels from 4 to 8 and 32 decreased the fit accuracy (by 10% and 50%, respectively). The additional channels had very little or no additional signal and mainly contributed noise.

Local single-channel approach

The Local-PseudoT1 approach combines the [Volz et al., 2012b] PseudoT1 regularizer with the [Mezer et al., 2013] local method. This approach is very efficient and works extremely well ($R^2 = 0.98$ and MAPE = 0.8% for the SOS M0, and $R^2 = 0.95$ and MAPE = 2% using the median of the channels). This result is almost identical to the T1 regularization with multichannel data (Fig. 6C).

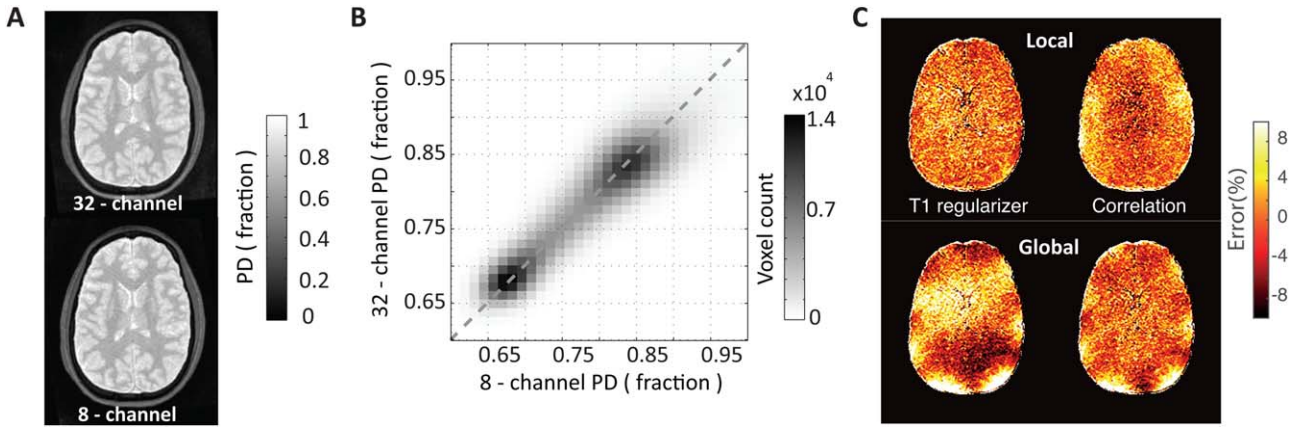

Figure 7.

Comparing in-vivo human brain PD measured with 8- and 32-channel coils. (**A**) An axial brain slice showing the PD map estimate by T1 local regularization in the same subject using two different RF coils (8-channel and 32-channel). (**B**) The 2D histogram compares the PD estimates measured using 8-channel (*x*-axis) and 32-channel (*y*-axis) coils ($R^2 = 0.89$). The gray scale measures the number of voxels. The two dark regions represent PD values from white matter (smaller) and gray matter (larger). (**C**) A comparison between four different PD- estimation methods. Spatial distribution of error is shown for two methods that use local PD fits (T1 local regularization, multichannel correlation [Mezer et al., 2013]) and two that use global fits (UNICORT [Volz et al., 2012a] and PseudoT1 [Volz et al., 2012b]). The error is calculated as the percent of the difference between the 8- and 32-channel PD images divided by the mean of the images.

Instrument Independence

We compared human-brain PD estimates calculated using two different coil arrays (Fig. 7); one with 8 channels and the other with 32 channels. The coil-sensitivity functions differ substantially between these arrays, due to both the size and spatial distribution of the coil elements. Thus, the comparison of the PD estimates from these two acquisitions evaluates the reliability in separating coil sensitivity from PD.

Using both single-channel and multichannel T1 local regularization (Fig. 7A,B), the two human- brain PD estimates agree for both gray and white matter ($R^2 = 0.89$). The reliability is similar using either multiple channels or combining the multiple channels into a single channel. This level of agreement is also similar to other recently published methods' multicoil correlation [Mezer et al., 2013] 0.88, UNICORT [Volz et al., 2012a] 0.87, and PseudoT1 [Volz et al., 2012b] 0.82. Importantly, the residual error size is similar across methods (MAPE ~5%), but the spatial distribution is only uniformly distributed for the T1 local regularization (Fig. 7C).

In-Vivo PD Mapping

We estimated the PD values of 15 subjects (ages 19–38). We calculated the mean and standard deviation of PD values across the gray matter (0.83 ± 0.06) and white matter (0.74 ± 0.07). These values agree with estimates using other MRI methods [Abbas et al., 2014, 2015; Fatouros and Marmarou, 1999; Gelman et al., 2001; Mezer et al., 2013; Tofts, 2003; Volz et al., 2012a,b] and non-MRI invasive methods [Tofts, 2003].

DISCUSSION

PD is the most basic MRI measurement, representing the percentage of observed water protons, the source of the MRI signal, in each voxel. The complement of PD, the lipid and macromolecular tissue volume (MTV), is a fundamental measure of human brain tissue.

Calibration and computational methods are required to separate the PD signal from the coil sensitivity. In modern practice, clinical magnets and coils are not calibrated, so that RF excitation and receive are major sources of unwanted variation [Tofts, 2003].

In this work, we describe several computational strategies to separate PD from coil sensitivity, *G* for spoiled-GE. We show that assuming smoothness of *G* is necessary but not sufficient; additional information and image-processing algorithms are needed. We compare different approaches that provide more information about *G*.

We find that multichannel data increases the robustness of the fit but also significantly increases the size of the acquired data and extends the computation time. We support the view of [Volz et al., 2012b] that T1 information has great benefit for two reasons: T1 regularization uses a biophysical prior that is more specific than general mathematical assumptions, and T1 regularization includes high-resolution spatial measurements that are important for

controlling overfitting of the smooth coil sensitivity *G*. Finally, global methods are less stable than local estimation methods.

Analyzing the local T1 regularization, we find that the additional advantage of multichannel data was minimal. Importantly, the Local-PseudoT1 method is the only one that has both high accuracy and uses a single channel for fitting. The ability to solve with single-channel data dramatically speeds up the analysis and renders the algorithm compatible with data that exist in many centers. Furthermore, multichannel approaches may be limited when each channel's data is under sampled due to accelerated acquisition with parallel imaging approaches. It is still advisable to test different methods, particularly in cases when there is reason to believe that T1 might not be a good predictor of PD (for example, when Gadolinium affects the signal), in pathological cases or when there is a significant amount of noise in the T1 estimates.

We confirm previous results showing a global correlation between 1/PD and 1/T1 in gray and white matter. The linear parameters [Eq. (4)] in this study are close to the typical white-matter values of $\gamma \sim 0.5$ s – 1 and $\delta \sim 0.8$ in previous studies [Abbas et al., 2015; Fatouros and Marmarou, 1999; Gelman et al., 2001; Mezer et al., 2013; Tofts, 2003]. Extending our previous work, we find that these parameters vary across the brain; furthermore, we also find significant local variation in the regression-line parameters. The PD value measures the lipid and macromolecular tissue volume (MTV = 1 − PD), while the T1 value measure depends on both the volume and composition of those lipids and macromolecules. We suggest that there is valuable information in describing the local values in individual subjects. Interestingly, a recent study proposed using a constant set of parameters relating T1 and PD to simultaneously map RF excite and sensitivity inhomogeneities [Baudrexel et al., 2016]. Further work is needed to determine the loss of accuracy caused by ignoring local parameter differences or parameter differences between individuals.

Parallel Imaging

Coil-sensitivity maps are an important part of parallel imaging. In general, improving the accuracy of coil-sensitivity maps enables greater acceleration [Blaimer et al., 2004; Larkman and Nunes, 2007].

Several earlier studies developed a framework for enhancing parallel imaging by jointly estimating coil sensitivity and spin density [Ying and Sheng, 2007] using regularization approaches [Liang et al., 2002; Lin et al., 2004; Uecker et al., 2008]. The current work uses many of the methods of parallel imaging, such as regularization, to estimate coil-sensitivity maps and avoid overfitting. We find that these approaches increase the accuracy of PD mapping, and that the biophysical method is the most reliable and better than the current practice of estimating coil sensitivity.

The methods we describe are not applicable to all cases of parallel imaging, but there may be some application in cases in which T1 maps are acquired. T1 regularization and the methods developed here for combining data from local ($\sim$2.7 cm^3), overlapping, low-order-polynomial coil-sensitivity maps may provide accurate maps that will be useful for accelerating SENSE applications. Extending this work into parallel imaging requires further research incorporating the phase component.

Limitations

T1-regularization assumes a linear PD-T1 relationship. In principle, large deviations from this relationship might yield poor PD estimates. The simulations show, however, that the linear relationship that is part of the T1 regularizer does not need to be very accurate (see Supporting Information Appendix 2).

The simulations and analyses neglect T2* effects on the MRI signal because the measurements are acquired using a short TE ($\sim$2 ms). Some T2* may be present in the data and affect the PD estimates [Abbas et al., 2014, 2015; Volz et al., 2012a]. There is a possibility of taking advantage of T2 or T2* mapping with multiple TEs that can be regressed to TE = 0 [Neeb et al., 2006; Whittall et al., 1997].

CONCLUSIONS

Using single channel data to separate PD from coil sensitivity is an ill-posed problem. One approach to solving the problem is to use multi-channel measurements. We show that in the absence of noise, a unique and accurate separation can be obtained by imposing the mathematical constraint of smooth coil sensitivities (low-order polynomials). In the presence of noise, however, such solutions are sensitive to overfitting and additional regularization terms are necessary. Using a biophysical prior about the T1-PD relationship yields more accurate estimates than any of the mathematical regularizations we tried. Further, applying the T1-PD constraint locally, and then joining the local solutions, provides particularly accurate PD estimates. Combining local T1-PD regularization with the assumption of smooth coil functions yields robust and accurate separation of PD from coil gain. Together these regularizers separate PD and coil gain effectively even when using conventional single channel data.

REFERENCES

Abbas Z, Gras V, Möllenhoff K, Keil F, Oros-Peusquens A-M, Shah NJ (2014): Analysis of proton-density bias corrections based on T1 measurement for robust quantification of water content in the brain at 3 Tesla. Magn Reson Med 72:1735–1745.

Abbas Z, Gras V, Möllenhoff K, Oros-Peusquens A-M, Shah NJ (2015): Quantitative water content mapping at clinically relevant field strengths: A comparative study at 1.5T and Neuroimage 106:404–413.

Avants B, Gee JC (2004): Geodesic estimation for large deformation anatomical shape averaging and interpolation. Neuroimage 23(Suppl 1):S139–S150.

Barral JK, Gudmundson E, Stikov N, Etezadi-Amoli M, Stoica P, Nishimura DG (2010): A robust methodology for in vivo T1 mapping. Magn Reson Med 64:1057–1067.

Baudrexel S, Reitz SC, Hof S, Gracien R-M, Fleischer V, Zimmermann H, Droby A, Klein JC, Deichmann R (2016): Quantitative T_1 and proton density mapping with direct calculation of radiofrequency coil transmit and receive profiles from two-point variable flip angle data. NMR Biomed 29:349–360.

Blaimer M, Breuer F, Mueller M, Heidemann RM, Griswold MA, Jakob PM (2004): SMASH, SENSE, PILS, GRAPPA: How to choose the optimal method. Top Magn Reson Imaging 15:223–236.

Chang L-C, Koay CG, Basser PJ, Pierpaoli C (2008): Linear least-squares method for unbiased estimation of T1 from SPGR signals. Magn Reson Med 60:496–501.

Deoni SCL, Rutt BK, Peters TM (2003): Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state. Magn Reson Med 49:515–526.

Fatouros PP, Marmarou A (1999): Use of magnetic resonance imaging for in vivo measurements of water content in human brain: Method and normal values. J Neurosurg 90:109–115.

Fram EK, Herfkens RJ, Johnson GA, Glover GH, Karis JP, Shimakawa A, Perkins TG, Pelc NJ (1987): Rapid calculation of T1 using variable flip angle gradient refocused imaging. Magn Reson Imaging 5:201–208.

Gelman N, Ewing JR, Gorell JM, Spickler EM, Solomon EG (2001): Interregional variation of longitudinal relaxation rates in human brain at 3.0 T: Relation to estimated iron and water contents. Magn Reson Med 45:71–79.

Hoge WS, Brooks DH, Madore B, Kyriakos WE (2005): A Tour of Accelerated Prallel MR Imaging from a Linear Systems Perspective. Concepts Magn Reson 27A:17–37.

Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM (2012): FSL. Neuroimage 62:782–790.

Kim SG, Hu X, Uğurbil K (1994): Accurate T1 determination from inversion recovery images: Application to human brain at 4 Tesla. Magn Reson Med 31:445–449.

Kucharczyk W, Macdonald PM, Stanisz GJ, Henkelman RM (1994): Relaxivity and magnetization transfer of white matter lipids at MR imaging: Importance of cerebrosides and pH. Radiology 192:521–529.

Larkman DJ, Nunes RG (2007): Parallel magnetic resonance imaging. Phys Med Biol 52:R15–R55.

Liang Z-PLZ-P, Bammer R, Ji J, Pelc NJ, Glover GH (2002): Improved image reconstruction from sensitivity-encoded data by wavelet denoising and Tokhonov regularization. In: 5th IEEE EMBS International Summer School Biomedical Imaging, 2002.

Lin F-H, Kwong KK, Belliveau JW, Wald LL (2004): Parallel imaging reconstruction using automatic regularization. Magn Reson Med 51:559–567.

Mezer A, Yeatman JD, Stikov N, Kay KN, Cho N-J, Dougherty RF, Perry ML, Parvizi J, Hua LH, Butts-Pauly K, Wandell BA (2013): Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. Nat Med 19:1667–1672.

Neeb H, Zilles K, Shah NJ (2006): A new method for fast quantitative mapping of absolute water content in vivo. Neuroimage 31:1156–1168.

Noterdaeme O, Anderson M, Gleeson F, Brady SM (2009): Intensity correction with a pair of spoiled gradient recalled echo images. Phys Med Biol 54:3473–3489.

Stüber C, Morawski M, Schäfer A, Labadie C, Wähnert M, Leuze C, Streicher M, Barapatre N, Reimann K, Geyer S, Spemann D, Turner R (2014): Myelin and iron concentration in the human brain: A quantitative study of MRI contrast. Neuroimage 93: 95–106.

Tofts P (2004): Quantitative MRI of the brain: Measuring changes caused by disease. J Neurol Neurosurg Psychiatry 75:1511.

Uecker M, Hohage T, Block KT, Frahm J (2008): Image reconstruction by regularized nonlinear inversion–joint estimation of coil sensitivities and image content. Magn Reson Med 60:674–682.

Uecker M, Lai P, Murphy MJ, Virtue P, Elad M, Pauly JM, Vasanawala SS, Lustig M (2014): ESPIRiT—An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA. Magn Reson Med 71:990–1001.

Volz S, Nöth U, Deichmann R (2012a): Correction of systematic errors in quantitative proton density mapping. Magn Reson Med 68:74–85.

Volz S, Nöth U, Jurcoane A, Ziemann U, Hattingen E, Deichmann R (2012b): Quantitative proton density mapping: Correcting the receiver sensitivity bias via pseudo proton densities. Neuroimage 63:540–552.

Weiskopf N, Lutti A, Helms G, Novak M, Ashburner J, Hutton C (2011): Unified segmentation based correction of R1 brain maps for RF transmit field inhomogeneities (UNICORT). Neuroimage 54:2116–2124.

Whittall KP, MacKay AL, Graeb DA, Nugent RA, Li DKB, Paty DW (1997): In vivo measurement of T2 distributions and water contents in normal human brain. Magn Reson Med 37:34–43.

Ying L, Sheng J (2007): Joint image reconstruction and sensitivity estimation in SENSE (JSENSE). Magn Reson Med 57:1196–1202.