The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/203387

Please be advised that this information was generated on 2020-11-07 and may be subject to change.
Original contribution

Repeatability and reproducibility of longitudinal relaxation rate in 12 small-animal MRI systems

John C. Waterton⁎, Catherine D.G. Hines, Paul D. Hockings, Iina Laitinen, Sabina Ziemian, Simon Campbell, Michael Gottschalk, Claudia Green, Michael Haase, Katja Hassemer, Hans-Paul Juretschke, Sascha Koehler, William Lloyd, Yanping Luo, Irma Mahmutovic Persson, James P.B. O’Connor, Lars E. Olsson, Kashmira Pindoria, Jurgen E. Schneider, Steven Sourbron, Denise Steinmann, Klaus Strobel, Sirisha Tadimalla, Irvin Teh, Andor Veltien, Xiaomeng Zhang, Gunnar Schütz

⁎ Corresponding author at: Bioxydyn Ltd, Manchester Science Park, Rutherford House, Pencroft Way MANCHESTER M15 6SZ, United Kingdom.

Abbreviations: Av, Avance; B₀, applied magnetic field; CoV, coefficient of variation; DNE, dynamic no enhancement; FISP, fast steady-state free-precession; ICH GCP, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use Harmonised Tripartite Guideline for Good Clinical Practice; PV, ParaVision; T₁, longitudinal relaxation rate; r₁, longitudinal relaxivity; rms, root-mean-square; SNR, signal-to-noise ratio; Tₛ, longitudinal relaxation time; T₁w, T₁-weighted; T₂, transverse relaxation time

https://doi.org/10.1016/j.mri.2019.03.008

Received 5 July 2018; Received in revised form 29 January 2019; Accepted 8 March 2019
1. Introduction

Many useful MR biomarkers derive from measurements of the water proton longitudinal relaxation time, $T_1$, or alternatively the relaxation rate $R_1 = 1/T_1$. Errors in $R_1$ [1] are common, will propagate, and may damage the reproducibility and accuracy of the resulting MR biomarkers. Although considerable effort has been devoted to measuring and assuring the accuracy of $R_1$ in clinical MR [2–5] systems, there is little evidence for the cross-site reproducibility of $R_1$ measurements in MR systems designed for small-animal research. The lack of standardisation in preclinical imaging has been recognised as an important problem [6,7] which in the worst case could invalidate the findings from animal studies, or confound meta-analyses and translation.

Reproducibility in a valid phantom is an important and ethical prerequisite for reproducible values in vivo. Poor technical validation has been a major impediment to clinical translation of MR biomarkers [8]. An ideal $R_1$ phantom should be traceable [2]; resist biological, chemical and physical deterioration; perform effectively over a range of temperatures convenient and relevant for the users; cover the parameter range expected in subsequent studies; not exhibit physiologically unrepresentative MR characteristics such as radiation damping, convection, unphysiologic $T_2$, excessive self-diffusion, off-resonance chemical shifts, standing waves, or abrupt boundaries; interrogate the entire volume subsequently to be occupied by body parts being imaged; have dimensions suitable for the subject subsequently to be imaged (in this case rats and mice); be convenient for the intended users; and be cost-effective for the intended users. To meet these criteria, nickel agarose phantoms following the design of Christofferson et al. [9] were used.

Two distinct general approaches to MR standardisation have previously been employed. In the first, which we term “centrally-led”, a central organisation, often independent of the participating sites, is accountable for overall measurement accuracy and reproducibility. They mandate the phantom and acquisition protocol and analyse centrally. They may perform set-up and training at each participating site, instruct sites to repeat aberrant measurements, or even expel sites who cannot achieve the required accuracy. Centrally-led standardisation is common in clinical trials performed to ICH GCP [10,11], or where the MR measurement is regulated as a companion diagnostic [12]. In the second approach, which we term “institution-led”, each investigator is accountable for measurement accuracy in their own centre. They are responsible for their own acquisition and analysis, and for compliance with any guidelines for their chosen phantom. “Institution-led” standardisation is common in academic research and in single-centre studies. Although we expect “centrally-led” standardisation to provide better reproducibility than “institution-led” standardisation, in this work we modelled “institution-led” standardisation as this is more representative of practice in preclinical MR. The study was performed within an international consortium of imaging centres participating in the validation of imaging biomarkers [13], and developing reliable preclinical MR assays which would give comparable results in different laboratories. The aim of this work was to assess the repeatability and reproducibility of $R_1$ in a realistic rodent MR protocol. Simple simulations were performed in order to compare the likely propagation of reproducibility errors into a broad range of $R_1$-derived MR biomarkers.

2. Materials and methods

2.1. Preclinical phantom

Batches of 2% agarose with nickel chloride concentrations respectively of 0.50, 1.04, 2.02, 4.08 and 8.05 mM, with 0.05% sodium azide, were prepared centrally in Berlin and used to create identical phantoms (Supplementary Fig. S3.1) which were distributed to the participating laboratories. The phantoms were prepared and authenticated (supplementary material S3) in July 2017, shipped in August 2017, and the measurements were performed between December 2017 and February 2018.

2.2. MR methods

Thirteen centres involved in an international consortium for the validation of imaging biomarkers for drug safety assessment [13] were invited to participate. Where centres had access to more than one MR system, they were invited to submit data from multiple MR systems. Eleven centres agreed to participate, one of which (G) provided data from two different magnets (G1 and G2): in the analyses, G1 and G2 were treated as if from two different centres. Details of the 12 MR systems are given in Table 1. Eleven of the 12 MR systems (all except B) were in laboratories which regularly and routinely measure MR biomarkers in rodents, intending to translate their findings to create diagnostics or therapeutics to improve human health. Although the use of any particular manufacturer’s equipment was not mandated, all participating centres elected to employ Bruker Avance/ParaVision systems. An “institution-led” approach to standardisation was adopted. Pilot studies were performed only in centres B and G. No site training was performed, no quality control was imposed, nor were sites permitted to repeat their measurements to eliminate apparent outliers. Region-of-
interest (RoI) definition and T1 calculation were performed locally. Centres were asked to measure R1 by saturation recovery using a standard RARE sequence. (Additional measurements using an investigational fast steady-state free-precession (FISP) sequence designed for the consortium’s in vivo needs will be reported elsewhere). In an attempt to provide temperature stability and minimise susceptibility artefacts, each phantom was embedded in a cucumber (Supplementary Figs. S3.2 and S3.3). Centres were instructed to “allow the five cumbered phantoms to come to thermal equilibrium in the magnet before and after each acquisition. The entire protocol was run in each centre on two separate days, mean 2.7 days apart (range 1–13).

In ParaVision, the standard RARE T1 saturation-recovery measurement method “1 map RARE protocol” (Rat/Head/Relaxometry) was invoked. All images were coronal with 58 × 58 mm field of view, 128 × 128 matrix, with a π/2 for 1.16 mm slice selection followed by a π train with RARE factor 8, effective echo time 30 ms, echo spacing 7.6 ms. Signal averaging was not employed and 5 dummy scans were used. Saturation recovery experiments used repetition times TR of 5500, 2000, 1200, 750, 500, 300, 200 and 100 ms giving a scan time of 169 s, not including the dummy scans. Next, a “dynamic-no-enhancement” (DNE) stability series to simulate dynamic contrast-enhanced MRI was run for 5 min (approximately 34 images) with repeated acquisition using the same parameters but with TR fixed at 500 ms.

### Table 1

| Centre | B0/T | Spectrometer          | Gradient strength/mT m−1 (model) | Radiofrequency transmitter/receiver volume coil (i.d./mm) | Software |
|--------|------|-----------------------|-------------------------------|---------------------------------------------------------|----------|
| A      | 7T   | Pharmascan 70/16 US Av III | 375 (B-GA96)                 | Quadrature 300 MHz (38)                                 | PV 6.0   |
| B      | 3    | Biospec 3 T Av IIHD    | 900 (B-GA105S HP)            | Quadrature 128 MHz (60)                                 | PV 6.0.1 |
| C      | 7    | Biospec 70/20 US Av IIHD | 660 (B-GA125 HP)             | Quadrature 300 MHz (86)                                 | PV 6.0.1 |
| D      | 4.7  | Biospec 47/20 US Av IIHD | 660 (B-GA125 HP)             | Quadrature 200 MHz (72)                                 | PV 6.0.1 |
| E      | 7    | Biospec 70/30 US Av II  | 440 (B-GA125S)               | Single channel 300 MHz (72)                              | PV 6.0.1 |
| F      | 7T   | Biospec 70/20 Av I     | 400 (B-GA12)                | Single channel 300 MHz (72)                              | PV 5.1   |
| G1     | 7    | Biospec 70/30 US Av III | 300 (B-GA12)                | Quadrature 300 MHz (90)                                 | PV 6.0.1 |
| G2     | 4.7  | Biospec 47/40 Av III   | 200 (B-GA125S)              | Quadrature 200 MHz (90)                                 | PV 6.0.1 |
| H      | 4.7  | Pharmascan 47/16 Av III | 300 (B-GA96)                | Single channel 200 MHz (60)                              | PV 5.1   |
| J      | 4.7  | Biospec 47/40 US Av II  | 660 (B-GA125 HP)            | Quadrature 200 MHz (72)                                 | PV 6.0.1 |
| K      | 9.4T | Biospec 94/30 Av III   | 670 (B-GA 125 HP)           | Quadrature 400 MHz (87)                                 | PV 6.0.1 |
| L      | 11.7 | Biospec 117/16 US Av III | 750 (B-GA 96)             | Quadrature 500 MHz (72)                                 | PV 6.0.1 |

Concentration of relaxive substance and in some instances to absolute biomarker value, while coefficients of variation may be more relevant when biomarker change is considered. Post-hoc tests of significance were made for “effect of day” using Student’s t-test, and for “effect of RoI position” by analysis of variance. No correction for multiple comparisons was made but p < 0.01 was considered significant. For each centre, weighted mean R1 values were calculated for each of the five phantoms:

\[
R_1 = \sum_{d=1,2} \sum_{RoI=0,X,Z} \left( \sum_{i=0,1} \left( \sum_{d=1,2} \sum_{RoI=0,X,Z} w_{d,Roi} \times R_{i,d,RoI} \right) \right) \left( \sum_{d=1,2} \sum_{RoI=0,X,Z} w_{d,Roi} \right)
\]

where \(R_{i,d,RoI}\) are the R1 values for each of the five RoIs and \(w_{d,Roi}\) are the corresponding weights, derived from the \(T_1\) fit in ParaVision:

\[
w = (\text{fitted slope}/\text{SD of fit})^2
\]

These weighted mean R1 values were then used to obtain relaxivities by linear regression:

\[
R_1/s = \frac{r_{i,B0} \times [Ni] + R_{i,[Ni]}=0,B0} + \epsilon
\]

where \(r_{i,B0}/s \cdot M^{-1}\) is the longitudinal relaxivity of aqueous Ni2+ in 2% agarose at field B0, \(R_{i,[Ni]}=0\), \(B0/s \cdot M^{-1}\) is the longitudinal relaxation rate of 2% agarose at field B0, and \(\epsilon\) is a normally-distributed error term assumed to subsume inter alia any temperature effects.

### 2.4. Cross-validation

Our “institution-led” study design required each centre to derive its own \(T_1\) values. Since centres elected to use the proprietary ParaVision software, a small supplementary study was also performed using an alternative analysis to verify values. Data from one centre were re-analysed. Centre A’s data were considered a good test set because they submitted data with both high and low fit errors. For each of the 10 RARE data sets (5 phantoms × 2 days), and for the same three RoIs used in the primary analysis, signal mean and standard deviation were retrieved for each TR value. \(R_1\) was calculated using “R” [14] using four expressions of the form:

\[
\text{model} = \ln(y[i, i])=(\text{Minf}-(\text{Minf} \times M0) \times \exp(-R1 \times (\text{TR} - 0.06))) \times \text{weights}(w[i, i])
\]

For three-parameter fits, Minf, M0 and R1 were fitted, while for two-parameter fits M0 was set to zero. For weighted fits, each RoI value

\[
\text{model} = \ln(y[i, i])=\left(\frac{\text{Minf}}{1 + \exp(-R1 \times (\text{TR} - 0.06))} \right) \times \text{weights}(w[i, i])
\]

For three-parameter fits, Minf, M0 and R1 were fitted, while for two-parameter fits M0 was set to zero. For weighted fits, each RoI value

\[
\text{model} = \ln(y[i, i])=(\text{Minf} \times \exp(-R1 \times (\text{TR} - 0.06))) \times \text{weights}(w[i, i])
\]
y was weighted by w, the inverse of the variance in y, while for un-weighted fits w was set to unity. For each of the 30 data sets, each of the four estimates of $R_1$ from “R”, $R_1^{PV}$, was compared with the reciprocal $T_1$ from Paravision, $R_1^{PV}$. In each of the four cases:

$$\text{mean difference} = \frac{1}{30} \sum \frac{R_1^{PV} - R_1^{PV}}{(R_1^{PV} + R_1^{PV})/2} \times 100\%$$

(2)

2.5. Illustrative simulations

Error propagation associated with two standard deviations of $R_1$ reproducibility was estimated for a range of derived measurements and biomarkers, using representative relaxivities and other parameters from the literature. This is conservative as it does not fully eliminate repeatability error. Three general cases were considered: firstly, native $R_1$ (or $T_1$) used as a biomarker; secondly, concentration of endogenous or exogenous paramagnetic substance used as biomarker; and thirdly, biomarkers derived from compartmental models. For Dynamic Contrast-Enhanced (DCE) MRI, the error in precontrast $R_1$ was propagated into the biomarkers for four preclinical case-studies. Representative ‘true’ values of kinetic parameters, pre-contrast $R_1$ values, and appropriate tracer kinetic models were chosen from literature to estimate contrast agent concentration uptake in each tissue type. Simulation parameters are provided in Supplementary Material.

3. Results

Each centre was requested to submit 30 $R_1$ measurements (5 phantoms × 3 locations × 2 days), the results of 10 DNE runs (5 phantoms × 2 days), and the 10 associated temperature measurements (5 phantoms × 2 days). The quality of the exponential fits for the 8 TR values was generally good, although in 15/360 cases the fit error was worse than 5% (9 cases in centre G2, 3 cases in centre A and 3 cases in centre L) (see Fig. 1). All these outliers were included in the analysis and not eliminated. One centre (J) did not provide DNE or temperature measurements in a suitable format, so its results were omitted from any analyses that needed those data. For the other centres, temperatures were recorded to ± 0.1°C: the mean was 19.3°C (SD 1.3), the mean deviation in temperature between day 1 and day 2 was 0.65°C, and the worst deviation 5°C (centre B, 0.5 mM phantom).

3.1. Longitudinal relaxation rates and relaxivities

Fig. 1 depicts the individual $R_1$ data, and Table 2 provides mean values. Fig. 2 shows the field dependence of $r_1$ from this work, with additional data points added from the literature [3,9,15,16].

3.2. Repeatability, reproducibility and linearity

Table 3 shows repeatability and reproducibility. Day-to-day repeatability ranged from 0.025 s⁻¹ (centre D) to 0.097 s⁻¹ (centre A): day-to-day repeatability CoV ranged from 0.76% (centre F) to 5.48% (centre L). In exploratory analysis, the day-to-day repeatability of 2.34% was not markedly improved either if measurements were
restricted to the isocentre (2.22%), or if measurements with > 1 °C difference in temperature between day 1 and day 2 were excluded (2.05%). No evidence was seen for field dependence of repeatability. For day-to-day repeatability, 2 centres (B, L) showed a statistically reproducible, on a CoV basis, was the 0.5 mM phantom at 4.7 T (2.94%, N = 4 centres) or, on an absolute units basis, the 8 mM phantom at 7 T (0.065 s⁻¹, N = 5 centres). In exploratory analysis, reproducibility was not improved if measurements were restricted to the isocentre (all-RoIs

rms reproducibility was 0.031 s⁻¹ or 1.4% while isocentre rms reproducibility was 0.064 s⁻¹ or 1.6%).

A measure of the linearity of \( R_1 \) as a biomarker over the range 0.8–8 s⁻¹ was obtained from the relaxation Eq. (1): the rms standard error of \( R_1 \), was 0.6% (range 0.2% in centre B, to 1.7% in centre L, \( N = 12 \) centres).

3.3. Comparison of analysis algorithms

\( R_1 \) values for centre A derived from two-parameter fits performed in "R" and in Paravision were close: mean differences were 0.024% for an unweighted fit and 0.26% for a weighted fit. When three-parameter fits performed in "R" were compared with two-parameter fits performed Paravision, disagreement was greater: 1.67% for an unweighted fit and 1.74% for a weighted fit. Bland-Altman style plots are provided in Supplementary Fig. S5.

3.4. Illustrative propagation to irreproducibility in biomarker values

Illustrative between-centre irreproducibility expected from two standard deviations of the observed \( R_1 \) reproducibility for a range of derived measurements and biomarkers are given in Table 4.

For measurements of concentration of substance, the propagated irreproducibility naturally varies with relaxivity, while for "derived" biomarkers the propagated irreproducibilities were generally ≤10%.

4. Discussion

In this work we addressed the repeatability and reproducibility of \( R_1 \) in MR systems designed and employed for translational in vivo research. We prefer to work with \( R_1 \) rather than \( T_1 \), since from a metrology perspective [17], \( R_1 \) is a ratio variable while \( T_1 \) is merely an interval variable. No single method for measuring \( R_1 \) is optimal for all in vivo studies. The most accurate methods (e.g. inversion recovery with long TR and short TE readout) are neither fast nor efficient. In vivo studies involve complex tradeoffs between accuracy, speed, spatial resolution, field of view, need for fat suppression, sensitivity to inflow, sensitivity to motion artefact, biexponential decay, and other confounding behaviours of tissue magnetisation such as \( T_2 \) and magnetisation transfer. Moreover, even after a specific method is chosen, errors can be very sensitive to pulse sequence parameters such as choice of delays and nutation angles, spoiling and refocussing strategies, mis-set pulses and so on. In this study we elected to use a RARE saturation recovery technique covering the entire field of view, as this is fairly robust and efficient: our findings may not be directly translatable to other commonly used techniques such as Variable Flip Angle [1,18,19] or Look-Locker [1,20,21] which are vulnerable to different confounds, or even to other saturation-recovery techniques with different pulse sequence parameters.

4.1. Repeatability and reproducibility

Previous work in preclinical MR systems has addressed the between-centre reproducibility of apparent diffusion coefficients [22] and volumetrics [23], but there is little evidence on relaxation rates. Clinical MR systems are designed, maintained and operated under Medical Device regulations, but these engineering and regulatory constraints do not apply to preclinical systems, so their reproducibility might differ from clinical reproducibility.

Repeatability [24,25] (ISO 3534:2:3.3.5) refers to the similarity between measurements over a short interval made using the same test object in the same equipment operated by the same investigator. Repeatability is particularly important when the same MR biomarker is measured on successive occasions in the same human or animal, for example before and after treatment. Repeatability depends on signal-to-noise ratio and on factors such as motion artefact, for which phantoms
Propagation of errors using Table 3 reproducibility, with plausible or representative values for a range of important measurements and biomarkers. Actual error propagation varies widely between applications: the values here should therefore be regarded as indicative, but not as a substitute for a thorough analysis of error propagation in any particular setting.

| Measurement or biomarker | Reproducibility error propagated from 2SD of R1 | Notes |
|--------------------------|-----------------------------------------------|-------|
| Native R1 (or T1)        | 0.062s⁻¹                                   |       |
| Tissue temperature       | 1.6–4.6°C                                    |       |
| Contrast agents          |                                               |       |
| Small non-protein-bound agents e.g. gadodextrate, gadopentetate, gadobutrol, relaxivities 3–11s⁻¹mM⁻¹ [62–67]. | 6–21μM |       |
| Gadobutrol in plasma at 9.4T [65], relaxation 4.7s⁻¹mM⁻¹ | 13μM |       |
| Gadodextrate, relaxivity [68] 5–17s⁻¹mM⁻¹ | 4–12μM |       |
| Ferumoxytiron iron oxide nanoparticles, relaxivity [69] of 20h⁻¹(mM Fe)⁻¹ at 1.5 T, monodisperse particle weight of 750kDa [70] | 3μM (Fe) or 0.2nM (particles) |       |
| Investigational folate dendrimer contrast agent with relaxivity [57] 1646s⁻¹mM⁻¹ | 38nM |       |
| Other substances         |                                               |       |
| Deoxyhaemoglobin monomer, relaxivity [56] 0.008s⁻¹mM⁻¹ | 7.8μM |       |
| Tempol (investigational radioprotectant), relaxivity [71] 0.2s⁻¹mM⁻¹ | 0.3μM |       |
| Dissolved dioxygen, relaxivity [72] 0.1–0.3s⁻¹mM⁻¹ | 160–470mmHg |       |
| Derived biomarkers       |                                               |       |
| Transfer constant Ktrans for gadopentetate in rodent glioma, extended Tofts model [73–75] | 0.004min⁻¹ (8%) |       |
| Extracellular extravascular fraction ve in rodent glioma, extended Tofts model [73–75] | 0.024 (10%) |       |
| Plasma fraction vP in rodent glioma, extended Tofts model [73–75] | 0.0016 (10%) |       |
| Transfer constant Kd for gadodextrate in transient ischemia model, Patlak analysis [73,76,77] | 0.0002μg⁻¹S⁻¹ (5%) |       |
| Plasma fraction vp, transient ischemia model, Patlak analysis [73,76,77] | 0.0008 (5%) |       |
| Flow Fp, normal rodent lung, model-free deconvolution [73,78,79] | 0.03min⁻¹ (8%) |       |
| Plasma fraction vN, normal rodent lung, model-free deconvolution [73,78,79] | 0.04 (10%) |       |
| Normal hepatocyte transporter uptake rate constant Kt for gadoxetate, 2-compartment liver model [73,80–82] | 0.0013mM.s⁻¹ (4%) |       |
| Normal hepatocyte transporter efflux rate constant Kve for gadoxetate, 2-compartment liver model [73,80–82] | 0.0001s⁻¹ (2%) |       |
| Extracellular extravascular fraction ve, 2-compartment liver model [73,80–82] | 0.016 (7%) |       |
| Albumin concentration    | 24μM (−5%)                                   |       |
| Extracellular matrix Fixed Charge Density | 8nM (−4%) |       |

Notes:

a Published data [53–55] suggest temperature dependence of tissue R1 in the range 0.013–0.039 s⁻¹°C.

b This very high relaxivity is per dendrimer molecule, not per Gd.

c The physiologic range is up to 17.5g∙dL⁻¹ (11mM).

d The physiologic range is 0–100mmHg in normoxia, 0–600mmHg in hyperoxia, >1000mmHg with hyperbaric oxygen.

e See supplementary material

f Typically drops in Ktrans of > 20% are pharmacologically significant [59].

h A drop in Ktrans of > 20% is toxicologically significant [80].

i For an albumin concentration of around 500μM, based on Eq. (13) and parameters from Fig. 1 in [85]. The physiologic and pathophysiologic range is approximately 450–750μM.

k Using Eq. (3) and cartilage data from Fig. 2 in [86]. These authors state “…assuming a 10% decrease in T1 would be measurable…we would expect to be sensitive to a change in FCD from a normal of −0.2 to −0.16M, the sort of change one would expect to see relatively early in a degenerative process.”

\[ \text{Notes:} \]

\[ ^{n} \text{Published data [53–55] suggest temperature dependence of tissue } R_{1} \text{ in the range } 0.013–0.039 \text{ s}^{-1} \cdot \text{C.} \]

\[ ^{b} \text{This very high relaxivity is per dendrimer molecule, not per Gd.} \]

\[ ^{c} \text{The physiologic range is up to } 17.5 \text{ g} \cdot \text{dL}^{-1} (11 \text{mM}). \]

\[ ^{d} \text{The physiologic range is } 0–100 \text{ mmHg in normoxia, } 0–600 \text{ mmHg in hyperoxia, } >1000 \text{ mmHg with hyperbaric oxygen.} \]

\[ ^{e} \text{See the supplementary material} \]

\[ ^{f} \text{Typically drops in } K_{\text{trans}} \text{ of } > 20\% \text{ are pharmacologically significant [59].} \]

\[ ^{h} \text{A drop in } K_{\text{trans}} \text{ of } > 20\% \text{ is toxicologically significant [80].} \]

\[ ^{i} \text{For an albumin concentration of around } 500\mu\text{M, based on Eq. (13) and parameters from Fig. 1 in [85]. The physiologic and pathophysiologic range is approximately } 450–750\mu\text{M.} \]

\[ ^{k} \text{Using Eq. (3) and cartilage data from Fig. 2 in [86]. These authors state “…assuming a 10% decrease in } T_{1} \text{ would be measurable…we would expect to be sensitive to a change in FCD from a normal of } −0.2 \text{ to } −0.16\text{M, the sort of change one would expect to see relatively early in a degenerative process”}. \]

The relaxivity of Ni(H2O)6²⁺ arises because two of the 3d nickel electron spin states are almost independent of B₀, as the Zeeman splittings are dominated by spin-orbit coupling (zero field splitting) and not by the applied field B₀. Above 2 T, the Zeeman splittings increase linearly with B₀. The relaxivity occurs through proton-electron dipolar mechanisms, with the relevant spectral density being the longitudinal relaxation rate R₁,e of the nickel electrons [31]. At low B₀, R₁,e depends on fluctuations of the zero field splitting which are independent of B₀, and previous investigators, working at relatively low fields, reported little field dependence for nickel agarose water proton T₁ values [15]. However our data, taken together with previous work (Fig. 2), clearly show a modest increase in relaxivity over the range 0.1–11.7 T.

4.2. Implications for translational research

Reproducibility errors (same subject, same device) have previously been extensively studied. Good reproducibility in phantoms is a necessary, but not sufficient, condition for good repeatability in vivo, because phantoms seldom model physiologic variability. However reproducibility errors (between centres) are much less studied, but are critically important in translating from single-centre to multi-centre use. Since physiologic variability is largely absorbed in the repeatability error, phantoms can be very informative about reproducibility.

Water proton T₁ was arguably the first MR biomarker [32–36].
Native $R_1$ has been reported as a biomarker *inter alia* in cardiac diseases [37,38], liver diseases [35,39,40], neurology [41], oncology [34,42], in the placenta [43] and in the lung [44–46]. Clinically significant $R_1$ differences (Table 4) usually exceed the expected irreproducibility reported in Table 3. For example: in liver fibrosis 0.1–0.2 s−1 or 10–20% [35,39,40]; in manganese neurotoxicology 0.06 s−1 or 7% [47]; in chronic obstructive pulmonary disease 0.1 s−1 or 10% [44] were clinically significant. In preclinical tumour models, differences of 15–20% were biologically significant [34]. Notably, however, in myocardial fibrosis, differences as small as 0.02–0.03 s−1 or 2–3% may be clinically significant [48,49] and in multiple sclerosis normal-appearing white matter differences of 0.025 s−1 or 2% may be clinically significant [50], so translational animal studies of these conditions may require exceptional efforts to ensure $T_1$ measurements can be validated and qualified for decision-making for these specific indications.

A second class of imaging biomarkers attaches a specific interpretation of the observed longitudinal relaxation, for example in arterial spin labelling [51,52] or in MR thermometry [53–55]. Thirdly, $R_1$ is commonly used to determine the spatially resolved in vivo concentration of an exogenous or endogenous paramagnetic substance of known relaxation. Relaxivity can be field-, tissue- and temperature-dependent, and varies over many orders of magnitude between rela

...
