Magnetic Resonance Imaging 59 (2019) 53–60

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.elsevier.com/locate/mri

Original contribution

Analytical validation of single-kidney glomerular filtration rate and split renal function as measured with magnetic resonance renography

Susmita Basak[a,*], David L. Buckley[a], Constantina Chrysochou[b], Anita Banerji[a], Diana Vassallo[b], Aghogho Odudu[c], Philip A. Kalra[b], Steven P. Sourbron[a]

[a] Biomedical Imaging Sciences Department, LICAMM, University of Leeds, LIGHT Laboratories, Leeds LS2 9JT, UK
[b] Department of Renal Medicine, Salford Royal NHS Foundation Trust, Stott Lane, Salford M6 8HD, UK
[c] Division of Cardiovascular Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, 46 Grafton Street, Manchester M13 9NT, UK

1. Introduction

Glomerular filtration rate (GFR) is a key clinical index of renal function and the basis for the definition and staging of chronic kidney disease. In clinical practice GFR is commonly assessed via estimated GFR (eGFR), which is based on an empirical conversion of blood creatinine values. eGFR is cheap, safe and widely available, but also suffers from low accuracy. Repeated blood sampling of injected radio-isotopes provides a more accurate measurement, but is time consuming and limited by the use of ionising radiation [1]. In recent years a non-ionising approach using iohexol has become more widely available, but as all the above techniques it does not measure single-kidney GFR (SK-GFR) or split renal function (SRF) [2,3]. This can currently be only obtained using separate scintigraphy, or by nuclear medicine renography [1,4,5].

GFR, SK-GFR and SRF can be measured simultaneously using magnetic resonance renography (MRR), which can be integrated in a routine MRI exam. An accurate MRR-measurement of these biomarkers could therefore potentially impact on the management of patients that already receive contrast-enhanced MRI for other reasons, and require a measured GFR, SK-GFR or SRF. Studies where an MRR could be integrated in the sequence protocol include, for instance, patients with renovascular disease considered for intervention [6–9], functional urinary obstruction [10], live kidney donors and post-transplant renal failure [11]. In addition, MRR also provides measurements of perfusion, vascularity and tubular transit times that can provide additional diagnostic power [12,13].

In order for MRR to fulfill this potential and replace the current gold-standard for GFR and SRF measurement, convincing evidence is required that the bias and precision is adequate for clinical practice (analytical or technical validity) [14]. A number of validation studies in small patient populations (cohort population size < 30) [13,15–17] have demonstrated strong correlation between MRR-based values (MR-GFR) and radioisotope gold standards. A validation of MR-GFR against a reference iohexol-GFR in healthy volunteers achieved low bias but only moderate precision [18]. A large, multi-centre study involving 295 patients with urinary obstructions reported clinically acceptable equivalence between MR-SRF and radioisotope SRF in a smaller subsection of 118 patients with moderately dilated kidneys, but failed to do so for severely dilated kidneys [10]. This study did not determine SK-GFR and was limited to a pediatric population.

The aim of this study was to determine the bias and precision of MRR-based measurements of SK-GFR, SRF and GFR in a larger population. A retrospective analysis of 127 studies with paired MRR and radio-isotope measurements of SK-GFR was performed, in a population covering a wide range of SK-GFR values. Secondary objectives were to investigate whether image processing and model refinement can improve the bias and precision. We included both 1 T and 3 T data in the study and analysed them separately to assess the effect of improvements in technology over time.

2. Material and methods

2.1. Patients

Data were collected from previous studies at Salford Royal Hospital (Manchester, UK) where paired MRR and radioisotope-SK-GFR were available with at most one week separating the measurements (1, 4, 5). Three of those studies, RVD1, RVD2 and RVD3, involved patients with atherosclerotic renovascular disease (ARVD) deemed to be suitable for revascularization [6–9,19,20] and the fourth group, Diab1 involved patients with diabetic kidney disease [21]. These studies were approved by institutional review board and informed consent was obtained from all patients. Data were excluded if MRR source data were incomplete or the reference values were unavailable. In all cases GFR was measured using standard nuclear medicine techniques [1,4,5,22]. As described in [22], GFR was measured following the administration of 3 MBq ^{51}Cr-EDTA diluted to 10 mL in 0.1% w/v excess EDTA solution. Injections

* Corresponding author at: LIGHT Laboratories, University of Leeds, Leeds LS2 9JT, UK.
E-mail address: s.basak@leeds.ac.uk (S. Basak).

https://doi.org/10.1016/j.mri.2019.03.005
Received 26 November 2018; Received in revised form 1 March 2019; Accepted 4 March 2019
0730-725X/

Table 1
Demographics.

Group	Duration	No. of patients/ studies	Age (mean)	Age (range) (years)	Gender		Radio-isotope SK-GFR range (mL/min)	Radio-isotope GFR range (mL/min)	Radio-isotope SK-GFR mean (mL/min)	Radio-isotope GFR mean (mL/min)	Field strength (Tesla)
					M	F					
Diab1	2006–Jun 2008	14/17	66	53–82	10	4	13–75	32–135	36.04	72.09	3.0
RVD1	2007–Jan 2010	33/47	65	39–83	20	13	1–61	13–111	22.29	44.59	3.0
RVD2	2004–Jan 2006	16/20	70	61–76	11	5	1–52	8–81	19.67	37.63	3.0
RVD3	Aug 2000–Dec 2002	38/43	69	34–84	27	11	0.43–74	10–101	16.58	32.03	1.0
Whole cohort	–	101/127	67	34–84	68	33	0.43–75	8–135	21.84	43.14	–

SK-GFR: single kidney glomerular filtration rate, GFR: glomerular filtration rate.

were administered just prior to ^{99m}Tc-dimercaptosuccinic acid (DMSA). Venous blood samples (10 mL) were taken 2, 3 and 4 h after injection. Samples were centrifuged and counted at least 72 h later to allow the decay of ^{99m}Tc. GFR was calculated from the slow exponential of the bi-exponential plasma clearance curve. The individual kidney function was calculated by dividing the GFR according to the percentage of the uptake of ^{99m}Tc-DMSA on scintigraphy. The SK-GFR was expressed as a true individual value in mL/min and not mL/min per 1.73 m^2 as body surface area was calculated for each patient.

2.2. MR measurements

MRR was performed at 3 T (Achieva; Philips) using phased-array body coils in three studies (Diab1, RVD1, RVD2) and at 1 T (Magnetom Expert; Siemens) using a combination of quadrature body coil and spine coil in one study (RVD3). A free-breathing 3D spoiled gradient echo sequence [13,19] with coronal-oblique slabs was used for the measurements. 3 T sequence parameters were: temporal resolution 2.1 s, FOV 400 × 400 × 80 mm, voxel size 3.13 × 3.13 × 4.0 mm^3, TR/TE 5.1/0.9 ms, FA 17°, SENSE factor 2. 1 T sequence parameters were: temporal resolution 4.5 s, FOV 350 × 306 × 80 mm, voxel size 2.7 × 2.7 × 2.5 mm^3, TR/TE 5.4/2.2 ms, FA 20°. 0.025 mmol/kg Gd-DOTA was injected at a rate of 3 mL/s for 3 T and 0.05 mmol/kg Gd-DTPA was hand-injected for 1 T. At 3 T a precontrast inversion recovery sequence was run for T1-mapping with inversion times 80, 500, 1400, 2250, 3850 ms.

2.3. MR image post-processing

Data were transferred in DICOM for offline post-processing using in-house software PMI 0.4 written in IDL 6.4 [23,24]. The author who performed the post-processing, a physicist with one year experience in medical imaging, was blinded to the reference values and patient information. Processing was performed as described in [13] with added Python scripts for automated analysis. The method is replicated exactly and therefore details are not repeated here. The analysis codes (details given in Section 2.5) are published on GitHub as supplementary materials. Briefly, an arterial ROI was defined in the aorta between the bifurcations of renal and iliac arteries to minimise inflow effects. Whole kidney ROIs were defined on a map of the contrast agent distribution volume [12] using thresholding and connected components to identify kidney clusters. Data were modelled assuming contrast agent concentration is proportional to signal enhancement and follows a two-compartment filtration model. GFR was determined by adding up SK-GFR of both kidneys and SRF was the ratio of left kidney SK-GFR to GFR.

To address the secondary objectives, three previously proposed model refinements were applied for the 3 T subgroup where T1 inversion-recovery sequence was available. The reference model described above is referred to as the “linear” approach. The first refinement approach (“linear + delay”) corrected for arterial delay times in the model fit [25]. The second refinement (“non-linear 1”) determined concentrations by inverting the steady-state spoiled gradient echo signal model [26], using literature values for T1 in blood (1.628 s [27]) and tissue (1.142 s [28]). The analysis was repeated using kidney T1 values measured with the inversion-recovery sequence (“non-linear 2”). A fourth refinement used a patient-specific value for hematocrit to calculate arterial plasma concentrations for a subgroup where hematocrit values were available.

2.4. Statistical analysis

For the primary objectives, comparisons between MRR values and radioisotope references are visualised through scatter plots and Bland-Altman plots. The following measures were derived for each of SK-GFR, GFR and SRF: mean and standard deviation (SD) of the difference with

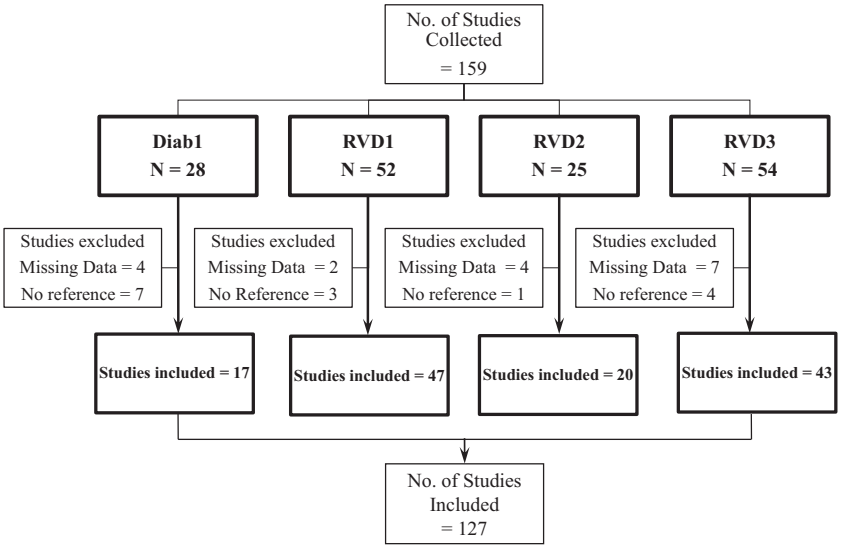

Fig. 1. Patient selection and distribution.

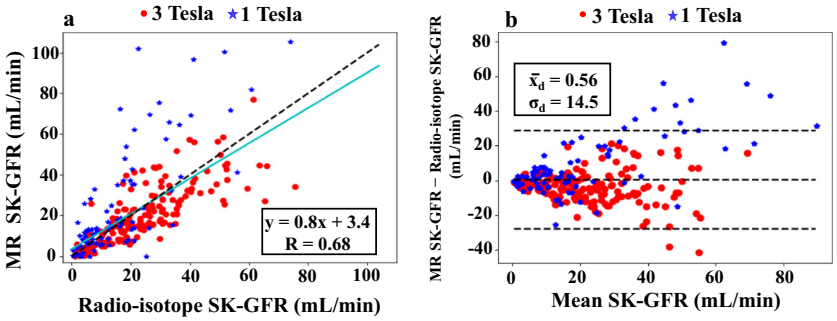

Fig. 2. Comparison of single-kidney glomerular filtration rate (SK-GFR) against reference method. (a) Regression analysis between MR SK-GFR and isotope SK-GFR. Solid (cyan) line is the regression line (equation and Pearson's coefficient (R) given in the figure) and dashed (black) line is the identity line. (b) Bland-Altman plot comparing MR SK-GFR and isotope SK-GFR. Dashed lines indicate the mean difference ($\bar{X}_d$) and 95% confidence intervals ($\bar{X}_d \pm 1.96\sigma_d$, σ_d = standard deviation of differences). Circles (red) and stars (blue) represent data for 3 Tesla and 1 Tesla subgroups, respectively.

the reference, 95% confidence interval (CI) for the difference (mean ± 1.96 SD), correlation coefficient, linear regression analysis and two-sided *t*-test for the difference. The bias in the measurements describes the central tendency of the data and it is determined by the mean of the difference of MRR values with the reference values [14]. On the other hand, the precision describes the dispersion in the data and it is determined by the SD of the difference of MRR values with the reference values [14].

As a benchmark of clinical utility, we calculated the percentage of values that were within 30% of the reference [29]. This aligns with guidance provided by the National Kidney Foundation's K/DOQI that any new method for GFR measurement has to do "substantially better" than the MDRD formula for estimated GFR (eGFR), which has found wide-spread clinical adoption based on evidence that 90% of the measurements were within 30% of the reference method.

For the secondary objective, mean, SD and p-value of the difference were calculated for 1 T and 3 T data separately, and for 3 T only for each of the 3 model refinements. The significance of the improvement of 3 T subgroup over 1 T subgroup was determined using two-sided *t*-tests. To test whether the model refinements affected bias and precision, the mean and variance of the difference with the reference were compared against the linear method with two-sided *t*-tests. All statistical computations were performed using the *stats* library in the *scipy* module of Python [30]. Statistical significance was defined as $p < 0.01$.

Apart from SRF, SK-GFR and GFR, the MRR method also produces other independent parameters, in particular renal blood flow, extracellular volume, filtration fraction and tubular mean transit time. These parameters were not evaluated in the current study as no reference measurements were available in this population, but the values are reported in the online tables.

2.5. Open access policy

All data and software used in this study are made freely available via GitHub (https://github.com/plaresmedima/Basak_et_al_2018) to allow secondary research and independent verification of the results. This includes signal-time curves for arterial input function (AIF) and kidney ROIs, Python scripts for automated processing, tables with results for each kidney, and a compiled PMI 0.4 version used for visualisation and analysis of the DICOM data. Anonymised DICOM data will be made freely available for secondary research, pending formal application and review by the project Steering Committee.

3. Results

3.1. Demographics

The patient characteristics are summarized in Table 1. In total 159 studies were collected, and 32 were excluded due to missing data, leaving 127 studies for analysis (Fig. 1). Repeated studies before and 4 months after renal artery revascularization were available for 26 patients. Hematocrit values were available for 52 studies.

3.2. Primary objective

The linear regression analysis and Bland-Altman plots for SK-GFR are shown in Fig. 2a and b, respectively. The results for GFR and SRF are in Fig. 3. Key statistics are summarized in Table 2, alongside comparable results from literature for reference. MRR results fell within

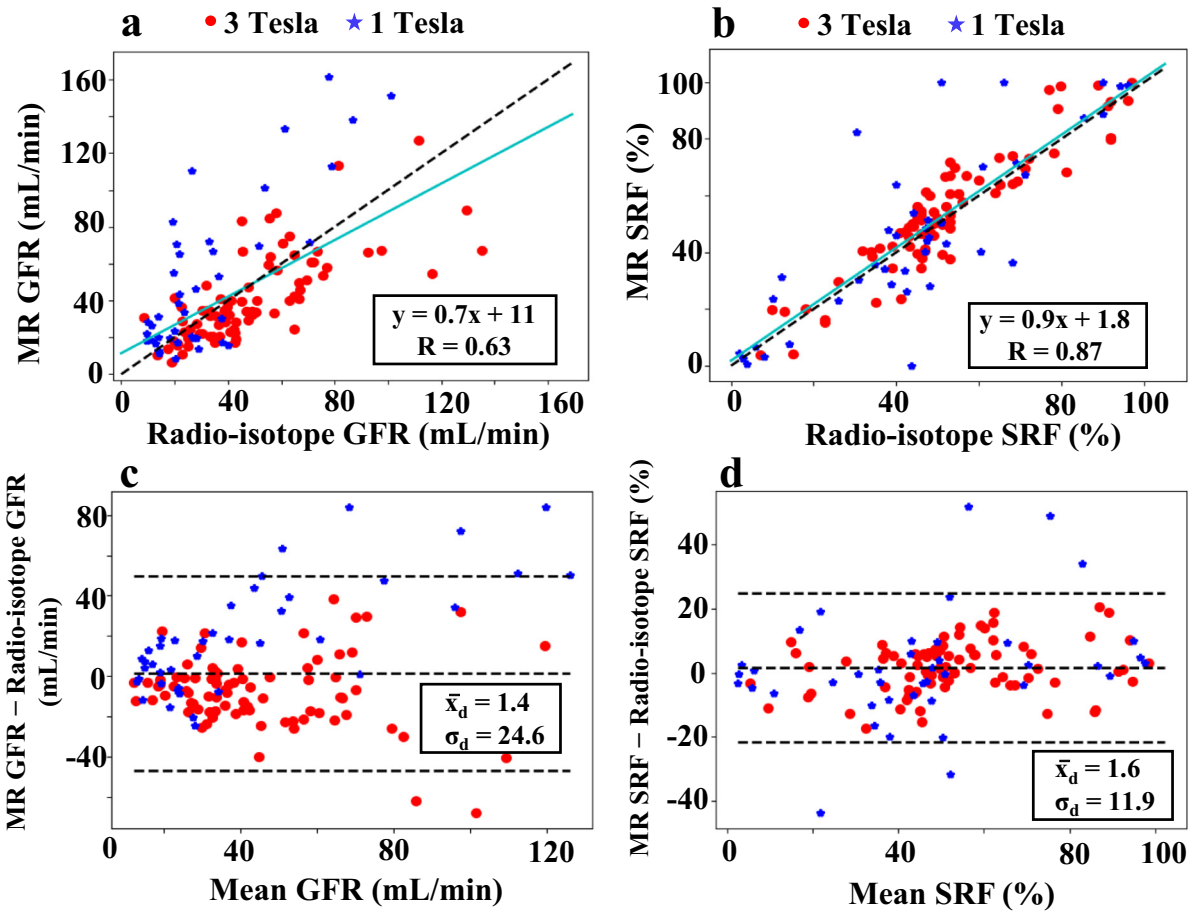

Fig. 3. Comparison of total glomerular filtration rate (GFR) and split renal function (SRF) against reference method. (a), (b) Regression analysis between MR GFR and isotope GFR and MR SRF and isotope SRF, respectively. Solid (cyan) line represents the regression line (equation and Pearson's coefficient (R) given in the figure) and dashed (black) line represents the identity line. (b), (d) Bland-Altman plots comparing MR total GFR and isotope total GFR and MR SRF and isotope SRF, respectively. Dashed lines indicate the mean difference ($\overline{X}_d$) and 95% confidence intervals ($\overline{X}_d \pm 1.96\sigma_d$, σ_d = standard deviation of differences). Circles (red) and stars (blue) represent data for 3 Tesla and 1 Tesla subgroups, respectively.

Table 2
Comparison of MRR against reference method. The table shows the results from this study (1 T and 3 T) and other patient studies from the literature for reference. GFR and SK-GFR values are given in mL/min and SRF values are given in %.

	Study	No. of studies	Slope	Intercept	R	Mean	SD	Mean diff	SD diff	95% CI	Within 30% of ref. values	p
SK-GFR	Our study (RVD and diabetes)	127	0.87	3.4 mL/min	0.68	22.4 mL/min	19.5 mL/min	0.56 mL/min	14.5 mL/min	−27.8 28.9 mL/min	41%	0.72
GFR	Our study		0.76	11 mL/min	0.63	44.5 mL/min	30.8 mL/min	1.39 mL/min	24.6 mL/min	−46.7 49.5 mL/min	45%	0.69
SRF	Our study		0.99	1.8%	0.87	0.51%	0.25%	1.60%	11.9%	−21.7% 24.9%	81%	0.59
SK-GFR	Lee 2007 (RVD) [15]	10	0.76	1.1 mL/min	0.84	–	–	−11.9 mL/min	13.4 mL/min	−38.2 14.5 mL/min	–	0.13
SK-GFR	Tipirneni-Sajja 2016 (Renal cancer) [17]	29	0.94	9.6 mL/min	0.87	78 mL/min	25 mL/min	−14.9 mL/min	11.8 mL/min	−39.0 13.0 mL/min	–	–
GFR	Vivier 2011 (Liver cirrhosis) [16]	20	–	–	0.88	–	–	−7.30 (Median diff) mL/min	12.8 (RMSE) mL/min	–	–	–
SRF	Claudon 2014 (Urinar obstruction) [10]	295	–	–	–	–	–	−2.00	14.4%	−30.2% 26.2%	–	–

MRR: magnetic resonance renography, SK-GFR: single kidney glomerular filtration rate, GFR: glomerular filtration rate, SRF: split renal function, R: correlation coefficient, CI: confidence interval, SD diff: standard deviation of differences, RMSE: root mean square error, p: p-value from two-sided *t*-test for equal means for difference of MR values from reference values.

Table 3
Comparison of later (3 T) versus earlier (1 T) technology. GFR values are given in mL/min and SRF values are given in %.

		Mean diff	SD diff
SK-GFR	3 T	−3.7 mL/min (p = 0.02)	9.72 mL/min
	1 T	9.06 mL/min (p = 0.008)	18.1 mL/min
	p-Value	< 0.001	< 0.001
GFR	3 T	−7.3 mL/min (p = 0.05)	17.7 mL/min
	1 T	18.9 mL/min (p = 0.01)	27.1 mL/min
	p-Value	< 0.001	< 0.001
SRF	3 T	1.7% (p = 0.59)	8.04%
	1 T	1.45% (p = 0.82)	17.2%
	p-Value	0.92	0.005

SK-GFR: single kidney glomerular filtration rate, GFR: glomerular filtration rate, SRF: split renal function, SD diff: standard deviation of differences. p-Value: p-value from two sided *t*-test for equal means (equal variances) for difference of MR values from reference values between 3 T and 1 T subgroup. p within bracket: p-value from two-sided *t*-test for equal means for difference of MR values from reference values.

30% of the reference values for 41% of SK-GFR, 45% of GFR and 81% of SRF. The main observation is that MRR shows low bias but poor precision. For example, the mean difference with the reference SK-GFR is just 0.56 mL/min, but the standard deviation is 14.5 mL/min.

3.3. Secondary objectives

The effect of advancement in technology is shown in Table 3. The results show that GFR and SK-GFR are less biased and more precise at 3 T than 1 T and reveal a systematic error at 1 T. The difference between 3 T and 1 T subgroup for both bias and precision of SK-GFR and GFR are significant as evident from the p-values. SRF has similar bias at 3 T and 1 T, but is more precise at 3 T. For the 3 T data alone, 50% of SK-GFR, 56% of GFR and 89% of SRF fell within 30% of the reference values.

The effect of model refinements is shown in Table 4. Refining the model did not improve the results – in particular the use of a non-linear signal model created a systematic error in SK-GFR and GFR, and further reduced the precision. In the 52 studies where a patient-specific hematocrit was available, hematocrit correction did not create any significant changes. The mean difference to reference SK-GFR increased by 3 mL/min and standard deviation of the difference increased by 1.6 mL/min.

For 3 T data with delay correction, 55% of SK-GFR, 65% of GFR and 90% of SRF were within 30% of the reference values. SRF meets the 90% acceptance limit set by the K/DOQI clinical practice guidelines for measurements of GFR performance.

3.4. Example cases

We provide two case studies, one at 1 T (Fig. 4) and one at 3 T (Fig. 5), to illustrate the data quality and typical issues. In all examples the model fits the data well. Fig. 4a shows a tortuous aorta near the edge of the MR volume. Both SK-GFR values are underestimated by an approximate factor of two, consistent with overestimation of the arterial concentrations due to signal increase at the edge of the slab. Fig. 5 shows a case with high precontrast signal intensities in the upper region of aorta due to inflow effects. This could lead to an error in the AIF, which was minimised by choosing the arterial ROI further down between the bifurcation of renal and iliac arteries in all cases. In this case the total GFR is very similar to the reference (64.5 mL/min vs 64.8 mL/min), but the SRF has a 12% difference between techniques (91% vs 79%). We don't have sufficient information to speculate the cause of this behavior but a likely candidate is the spatially inhomogeneous B1-effects, which could potentially be corrected with B1-mapping. We emphasize that these examples are purely anecdotal and no scientific inference regarding the trend of the data or performance of MRR should be interpreted from these examples.

4. Discussion

A reliable method to measure GFR, SK-GFR and SRF with MRR could have an immediate impact on the management of patients that currently require a separate radio-isotope measurement of these quantities. The aim of this study was to investigate whether the bias and precision of MRR is sufficient to justify a substitution of the current reference method with this new approach. MRR-based values were compared against a reference radio-isotope technique in a large retrospective cohort of 127 studies. The key observations are: (1) that MRR is low biased (small mean difference) but substantially less precise (large standard deviation of the difference) in all three parameters; (2) that the precision of MRR was worse in older (1 T) studies than more recent (3 T) studies; (3) that the precision of MRR could not be improved using a priori defined refinements of the analysis method.

This study is the largest validation study of MRR-based SK-GFR in adult patients. Considering only the results with more state-of-the-art hardware at 3 T with the optimal modeling approach “linear + delay”, the bias and precision improve on recent validation studies in patients with renovascular disease [15], urinary obstruction [10], liver cirrhosis [16] and renal cancer [17]. Claudon et al. [10] investigated 295 pediatric patients and reported a lower precision in SRF (SD 14% vs SD 7.8% in this study) and higher bias (2.0% mean difference vs 0.81% in this study). Lee et al. [15] and Tipirneni-Sajja et al. [17] measured SK-GFR error of 12 mL/min (SD 13) and 15 mL/min (SD 12), compared to 2.7 mL/min (SD 9.4) in this study. Vivier et al. do not provide directly comparable metrics, but their median GFR-difference of 7.3 mL/min is larger than the mean difference 5.3 mL/min measured in this study. One discrepancy with previous work is that Tipirneni-Sajja et al. [17] reported improved correlations after hematocrit correction, which was

Table 4
Results on the secondary objectives: effect of model refinements on 3 T subgroup. GFR values are given in mL/min and SRF values are given in percentage of total GFR.

	Linear		Linear + delay		Non-linear 1 (Literature value for T1)		Non-linear 2 (Measured T1)	
	Mean diff	SD diff	Mean diff	SD diff	Mean diff	SD diff	Mean diff	SD diff
SK-GFR	−3.7 mL/min	9.72 mL/min	−2.72 mL/min	9.39 mL/min	**10.1** mL/min	**18.3** mL/min	**4.99** mL/min	**14.08** mL/min
GFR	−7.3 mL/min	17.7 mL/min	−5.32 mL/min	17.2 mL/min	**20.5** mL/min	**32.4** mL/min	**8.49** mL/min	25.3 mL/min
SRF	1.7%	8.0%	0.81%	7.76%	3.11%	12.1%	2.54%	12.2%

SK-GFR: single kidney glomerular filtration rate, GFR: glomerular filtration rate, SRF: split renal function, SD diff: standard deviation of difference. Bold fonts: significantly different ($p < 0.01$) compared to the linear method, where p is the p-value from two sided t-test for equal means (equal variances) for difference of MR values from reference values between two methods.

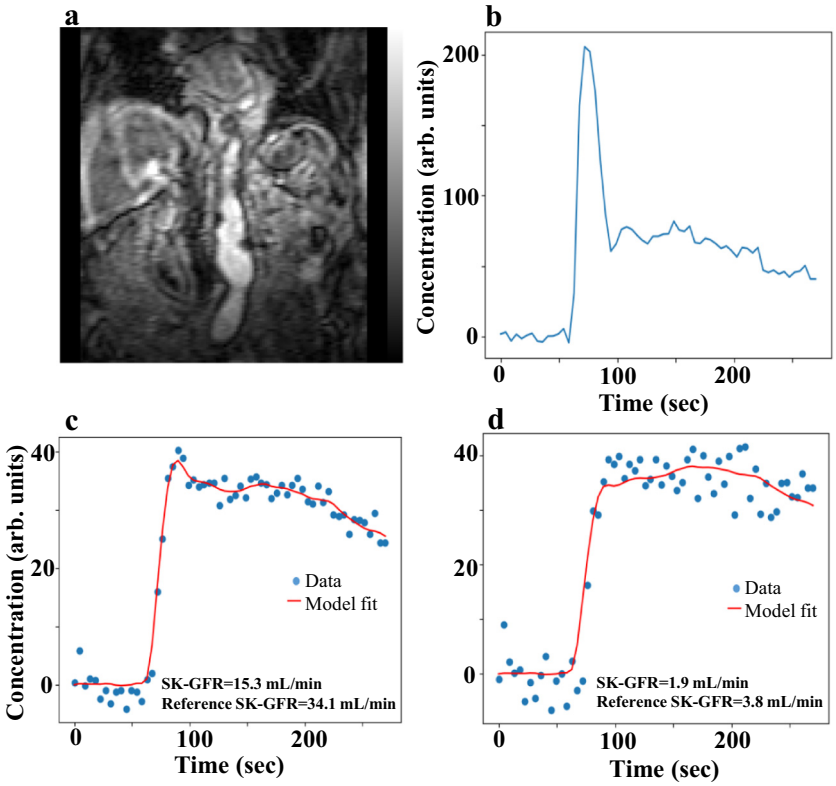

Fig. 4. Example case from 1 Tesla subgroup. (a) Dynamic contrast enhanced MR image (contrast agent Gd-DTPA) of a patient (age = 76 years) with tortuous aorta. The image slice is situated near the edge of the MR slab, (b) Arterial input function, (c), (d) concentration-time curve for left kidney and right kidney, respectively.

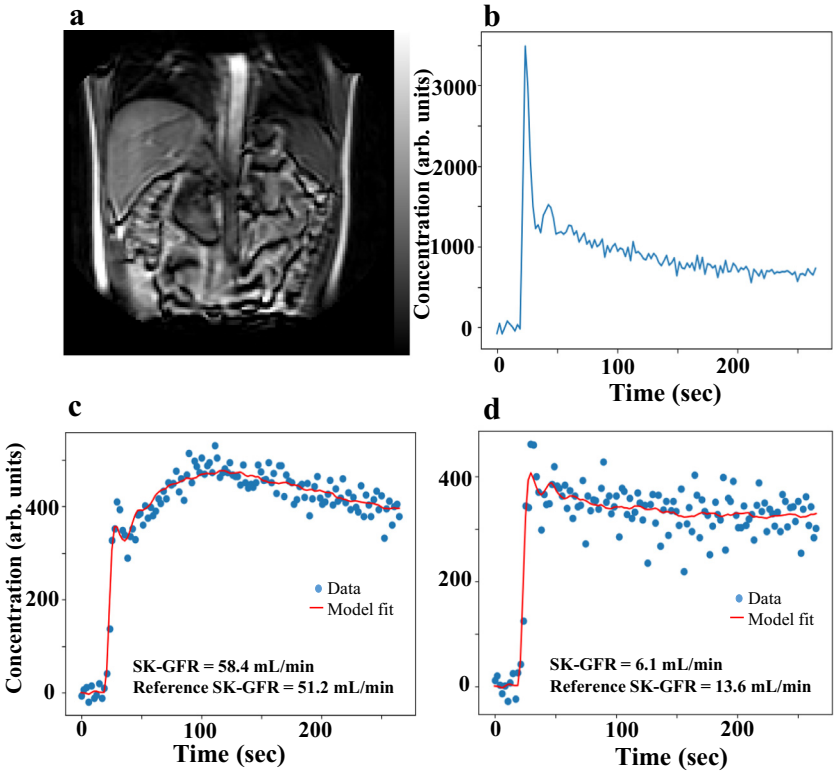

Fig. 5. Example case from 3 Tesla subgroup. (a) Pre-contrast image (contrast agent Gadoterate) showing the presence of large inflow for a patient (age = 52 years) , (b) Arterial input function, (c), (d) concentration-time curve for left kidney and right kidney, respectively.

not observed in this study.

The key question in terms of clinical utility is whether the observed precision is sufficient to justify a replacement of the reference method by the new MRR method in clinical practice. For GFR the guidance provided by the National Kidney Foundation's Kidney Disease Outcomes Quality Initiative (K/DOQI) [29] is that any new method has to do "substantially better" than the MDRD formula for estimated GFR (eGFR), which has found wide-spread clinical adoption based on evidence that 90% of the measurements were within 30% of the reference method. MRR falls significantly short of that: even when isolating the 3 T data, only 65% of GFR values are within 30% of the reference. We can therefore safely conclude that substantial technical improvements are required before MRR-based GFR can be adopted clinically.

For SRF the conclusion is more positive: at 3 T, 90% of the values are within 30% of the reference, and this fulfils the criterion that led to the acceptance of MDRD as a clinical tool. However, this is unlikely to be adequate for clinical practice: a 4% difference is considered clinically significant on ^{99m}Tc-DMSA, indicating the MRR-SRF is still far from being acceptable as a replacement for nuclear medicine investigations [31]. A more careful outcome analysis and health-economy modeling is required to support this conclusion, possibly combined with refinements in the method to reduce outliers. Nephrogenic systemic fibrosis (NSF) has been a contentious issue, but recent evidence and recommendations demonstrate that the risk with macrocyclic agents is negligible when current safety guidelines are adhered to [32].

We emphasize that the K/DOQI recommended guideline is being used as a useful benchmark and not as a target to confirm clinical adoption. MRR needs not only to achieve this benchmark but also be more accurate than creatinine-based eGFR. Moreover, this metric is useful because it is dimensionless and enables comparison between GFR and SRF which have different units and therefore, bias and precision are not directly comparable. For instance, we see consistently that this overall level of uncertainty is smaller in SRF than in SK-GFR, which is consistent with the fact that SRF is a ratio and therefore any scaling errors common to left and right kidney cancel out. This type of comparison is important to determine the main sources of error.

A limitation of this study is that it is retrospective and uses historical data acquired between 8 and 18 years ago. The acquisition methods therefore do not incorporate latest insights into MRR quantification or state-of-the-art acquisition and reconstruction methods. The older 1 T data were nevertheless included because the comparison of 1 T and 3 T data allowed us to determine the effect of technological advances. Since patients with two different types of kidney diseases have been included in the 3 T subgroup, the etiologies are largely, but not completely comparable between 1 T and 3 T subgroups. However, for the purposes of this study the key point is that the SK-GFR ranges for both populations are similar. We see that 1 T data is systematically higher than the 3 T data. The precise cause of this observation is unclear but there are clear improvements in bias and precision moving from 1 T to 3 T data demonstrates that MRR technology is on an upwards trajectory, and offers some confidence that further technical improvements will push the error safely below the benchmarks set by the National Kidney Foundation [29]. Also, the reproducibility of 3 T data, specifically for patients with very low values of SK-GFR, needs to be tested to confirm the improvement with technical advancements. Another limitation is the assumption of repeat measures as independent studies, which is a requirement in the statistical modeling. While this is not exactly satisfied for a repeat measurement, the 4 month time gap and the intervention between the two measurements does create some level of independence.

A key question for future development is to identify the main limiting factors to precision. The fact that SRF, a relative measure, is substantially more reliable than SK-GFR, indicates that global scaling errors are a major factor. This points to AIF-errors such as those caused by inflow effects (32, 33), which can be minimised in various ways, e.g. by extending the field of view to include the heart, and increasing slab thickness. The fact that non-linear signal analysis increased the bias points to false assumptions in the signal model, caused by for instance B1-effects [33] or imperfect spoiling [34], both of which can be minimised or corrected for by techniques such as B1-mapping or improved rf-spoiling [35].

The 3D data also suffered from significant intra- and inter-frame motion artefacts, which can be minimised using radial scanning, multislice 2D acquisitions, motion-compensated reconstruction [36] or motion correction [37]. Radial acquisitions in particular combined with novel reconstruction methods are promising in this respect, as they enable fast dynamic scanning as required by renography-type approaches, but without sacrificing much image quality compared to conventional breath-held T1-weighted sequences [38]. This is important as many clinical applications where SRF measurements are important (e.g. assessment of potential donors) require high-quality 3D images in arterial-, venous and excretory phases that cannot be compromised by replacing them with fast MRR data with significantly reduced image quality.

Most of these technical solutions cannot be implemented retrospectively and will require dedicated prospective studies. Since our study does not measure reproducibility or repeatability, but performs a comparison with the reference measurements, the error also potentially includes contributions from (subject-specific) systematic differences. The reproducibility error is expected to be lower as seen in Ref. [31]. The reproducibility of the MR data needs to be studied explicitly along with the improvement of the precision. In order to eliminate risk from such expensive and time-consuming clinical validation studies, it may be prudent to first evaluate possible approaches in-silico on computational phantoms of human body. The anonymised DICOM data from this study are made freely available to help inform these further developments.

5. Conclusions

In summary we conclude: (1) that MRR-based SK-GFR has low bias but is currently too imprecise with performance substantially poorer than the creatinine-based eGFR and requires further technical development before clinical adoption can be considered; (2) that MRR-based SRF measured at 3 T has similarly low bias and is more precise, but remains far removed from the precision of radio-isotope methods for SRF, and therefore not ready for adoption in clinical practice. More evidence from prospective clinical studies incorporating the latest improvements in MRI hardware and reconstruction are needed to arrive at more definitive conclusions.

Acknowledgements

This work was funded by Kidney Research UK [grant numbers RP55/2012 and RP24/1/2006] and Wellcome Trust [award number 071760].

References

[1] Rehling M, Moller ML, Lund JO, Jensen KB, Thamdrup B, Trap-Jensen J. 99mTc-DTPA gamma-camera renography: normal values and rapid determination of single-kidney glomerular filtration rate. Eur J Nucl Med 1985;11(1):1–6.
[2] Brandstrom E, Grzegorczyk A, Jacobsson L, Friberg P, Lindahl A, Aurell M. GFR measurement with iohexol and 51Cr-EDTA. A comparison of the two favoured GFR markers in Europe. Nephrol Dial Transplant 1998;13(5):1176–82.
[3] Soveri I, Berg UB, Bjork J, Elinder CG, Grubb A, Mejare I, et al. Measuring GFR: a systematic review. Am J Kidney Dis 2014;64(3):411–24.
[4] Petersen LJ, Petersen JR, Talleruphuus U, Moller ML, Ladefoged SD, Mehlsen J, et al. Glomerular filtration rate estimated from the uptake phase of 99mTc-DTPA renography in chronic renal failure. Nephrol Dial Transplant 1999;14(7):1673–8.
[5] Rehling M, Moller ML, Thamdrup B, Lund JO, Trap-Jensen J. Reliability of a 99mTc-DTPA gamma camera technique for determination of single kidney glomerular filtration rate. A comparison to plasma clearance of 51Cr-EDTA in one-kidney patients, using the renal clearance of inulin as a reference. Scand J Urol Nephrol 1986;20(1):57–62.

[6] Cheung CM, Chrysochou C, Shurrab AE, Buckley DL, Cowie A, Kalra PA. Effects of renal volume and single-kidney glomerular filtration rate on renal functional outcome in atherosclerotic renal artery stenosis. Nephrol Dial Transplant 2010;25(4):1133–40.
[7] Cheung CM, Shurrab AE, Buckley DL, Hegarty J, Middleton RJ, Mamtora H, et al. MR-derived renal morphology and renal function in patients with atherosclerotic renovascular disease. Kidney Int 2006;69(4):715–22.
[8] Chrysochou C, Buckley DL, Dark P, Cowie A, Kalra PA. Gadolinium-enhanced magnetic resonance imaging for renovascular disease and nephrogenic systemic fibrosis: critical review of the literature and UK experience. J Magn Reson Imaging 2009;29(4):887–94.
[9] Chrysochou C, Buckley DL, Kalra PA. Magnetic resonance imaging: advances in the investigation of atheromatous renovascular disease. J Nephrol 2008;21(4):468–77.
[10] Claudon M, Durand E, Grenier N, Prigent A, Balvay D, Chaumet-Riffaud P, et al. Chronic urinary obstruction: evaluation of dynamic contrast-enhanced MR urography for measurement of split renal function. Radiology 2014;273(3):801–12.
[11] Yamamoto A, Zhang JL, Rusinek H, Chandarana H, Vivier PH, Babb JS, et al. Quantitative evaluation of acute renal transplant dysfunction with low-dose three-dimensional MR renography. Radiology 2011;260(3):781–9.
[12] Dujardin M, Sourbron S, Luypaert R, Verbeelen D, Stadnik T. Quantification of renal perfusion and function on a voxel-by-voxel basis: a feasibility study. Magn Reson Med 2005;54(4):841–9.
[13] Lim SW, Chrysochou C, Buckley DL, Kalra PA, Sourbron SP. Prediction and assessment of responses to renal artery revascularization with dynamic contrast-enhanced magnetic resonance imaging: a pilot study. Am J Physiol Renal Physiol 2013;305(5):F672–8.
[14] Sullivan DC, Obuchowski NA, Kessler LG, Raunig DL, Gatsonis C, Huang EP, et al. Metrology standards for quantitative imaging biomarkers. Radiology 2015;277(3):813–25.
[15] Lee VS, Rusinek H, Bokacheva L, Huang AJ, Oesingmann N, Chen Q, et al. Renal function measurements from MR renography and a simplified multicompartmental model. Am J Physiol Renal Physiol 2007;292(5):F1548–59.
[16] Vivier PH, Storey P, Rusinek H, Zhang JL, Yamamoto A, Tantillo K, et al. Kidney function: glomerular filtration rate measurement with MR renography in patients with cirrhosis. Radiology 2011;259(2):462–70.
[17] Tipirneni-Sajja A, Loeffler RB, Oesingmann N, Bissler J, Song R, McCarville B, et al. Measurement of glomerular filtration rate by dynamic contrast-enhanced magnetic resonance imaging using a subject-specific two-compartment model. Physiol Rep 2016;4(7).
[18] Eikefjord E, Andersen E, Hodneland E, Hanson EA, Sourbron S, Svarstad E, et al. Dynamic contrast-enhanced MRI measurement of renal function in healthy participants. Acta Radiol 2017;58(6):748–57.
[19] Buckley DL, Shurrab AE, Cheung CM, Jones AP, Mamtora H, Kalra PA. Measurement of single kidney function using dynamic contrast-enhanced MRI: comparison of two models in human subjects. J Magn Reson Imaging 2006;24(5):1117–23.
[20] Chrysochou C, Power A, Shurrab AE, Husain S, Moser S, Lay J, et al. Low risk for nephrogenic systemic fibrosis in nondialysis patients who have chronic kidney disease and are investigated with gadolinium-enhanced magnetic resonance imaging. Clin J Am Soc Nephrol 2010;5(3):484–9.
[21] Rutter MK, Prais HR, Charlton-Menys V, Gittins M, Roberts C, Davies RR, et al. Protection Against Nephropathy in Diabetes with Atorvastatin (PANDA): a randomized double-blind placebo-controlled trial of high- vs. low-dose atorvastatin (1). Diabet Med 2011;28(1):100–8.
[22] Farmer CK, Cook GJ, Blake GM, Reidy J, Scoble JE. Individual kidney function in atherosclerotic nephropathy is not related to the presence of renal artery stenosis. Nephrol Dial Transplant 1999;14(12):2880–4.
[23] Attenberger UI, Sourbron SP, Schoenberg SO, Morelli J, Leiner T, Schoeppler GM, et al. Comprehensive MR evaluation of renal disease: added clinical value of quantified renal perfusion values over single MR angiography. J Magn Reson Imaging 2010;31(1):125–33.
[24] Sourbron SP, Michaely HJ, Reiser MF, Schoenberg SO. MRI-measurement of perfusion and glomerular filtration in the human kidney with a separable compartment model. Invest Radiol 2008;43(1):40–8.
[25] Sourbron SP, Buckley DL. Tracer kinetic modelling in MRI: estimating perfusion and capillary permeability. Phys Med Biol 2012;57(2):R1–33.
[26] Sourbron S. Technical aspects of MR perfusion. Eur J Radiol 2010;76(3):304–13.
[27] Lu H, Clingman C, Golay X, van Zijl PC. Determining the longitudinal relaxation time (T1) of blood at 3.0 Tesla. Magn Reson Med 2004;52(3):679–82.
[28] de Bazelaire CM, Duhamel GD, Rofsky NM, Alsop DC. MR imaging relaxation times of abdominal and pelvic tissues measured in vivo at 3.0 T: preliminary results. Radiology 2004;230(3):652–9.
[29] National Kidney F. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis 2002;39(2 Suppl 1):S1–266.
[30] Python: a dynamic, open source programming language URL Python Software Foundation; 2015https://www.python.org/2015.
[31] Gordon I, Evans K, Peters AM, Kelly J, Morales BN, Goldraich N, et al. The quantitation of 99Tcm-DMSA in paediatrics. Nucl Med Commun 1987;8(8):661–70.
[32] Khawaja AZ, Cassidy DB, Al Shakarchi J, McGrogan DG, Inston NG, Jones RG. Revisiting the risks of MRI with gadolinium based contrast agents-review of literature and guidelines. Insights Imaging 2015;6(5):553–8.
[33] Roberts C, Little R, Watson Y, Zhao S, Buckley DL, Parker GJ. The effect of blood inflow and B(1)-field inhomogeneity on measurement of the arterial input function in axial 3D spoiled gradient echo dynamic contrast-enhanced MRI. Magn Reson Med 2011;65(1):108–19.
[34] Garpebring A, Wirestam R, Ostlund N, Karlsson M. Effects of inflow and radiofrequency spoiling on the arterial input function in dynamic contrast-enhanced MRI: a combined phantom and simulation study. Magn Reson Med 2011;65(6):1670–9.
[35] Georgiou L, Wilson DJ, Sharma N, Perren TJ, Buckley DL. A functional form for a representative individual arterial input function measured from a population using high temporal resolution DCE MRI. Magn Reson Med 2018;81(3):1955–63.
[36] Feng L, Axel L, Chandarana H, Block KT, Sodickson DK, Otazo R. XD-GRASP: golden-angle radial MRI with reconstruction of extra motion-state dimensions using compressed sensing. Magn Reson Med 2016;75(2):775–88.
[37] Dimitra Flouri DL, Chrysochou Constantina, Kalra Philip, Sourbron Steven P. Motion correction for 3D free-breathing renal DCE-MRI using tracer kinetic model-driven registration. The ISMRM 25th Annual Meeting & Exhibition. Honolulu, HI, USA. 2017.
[38] Chandarana H, Block TK, Rosenkrantz AB, Lim RP, Kim D, Mossa DJ, et al. Free-breathing radial 3D fat-suppressed T1-weighted gradient echo sequence: a viable alternative for contrast-enhanced liver imaging in patients unable to suspend respiration. Invest Radiol 2011;46(10):648–53.