The Renal Resistive Index in Allografts: Is Sonographic Assessment Sufficiently Reproducible in a Routine Clinical Setting?

Reproducibility of the Renal Resistive Index

ZUSAMMENFASSUNG

Ziel Die Überprüfung der Reproduzierbarkeit des renalen Widerstandsindex (RRI) in der Alltagssituation.


Ergebnisse Die statistische Auswertung ergab eine Reproduzierbarkeit des parenchymalen RRI von 0,045 bei einer Grenze von 0,124. Der ICC zwischen den parenchymalen und hilären RRIs lag bei 0,852 bzw. 0,868. Der Kruskal-Wallis-Test zeigte einen signifikanten Unterschied in der Reproduzierbarkeit bei unterschiedlichen Ultraschallgeräten (p = 0,003). Die Korrelation der Varianz der RRIs mit der Tiefe des Nierentransplants im Patienten war ebenfalls signifikant (p = 0,001).

Schlussfolgerung Die Reproduzierbarkeit des RRI ist prinzipiell auch in der Alltagssituation gegeben, nimmt aber mit der Tiefe des Transplantats im Patienten und durch die Benutzung unterschiedlicher Ultraschallgeräte ab.

Kernaussagen:
- Der renale Widerstandsindex (RRI) in Nierentransplantaten ist ausreichend reproduzierbar.
- Der gemessene RRI wird vom verwendeten Ultraschallgerät beeinflusst.
- Die Reproduzierbarkeit des RRI nimmt mit der Transplantattiefe im Empfänger ab.

ABSTRACT

Purpose To assess the reproducibility of the renal resistive index (RRI) in a routine clinical setting.

Materials and Methods 22 patients with a kidney allograft and 19 physicians participated in our prospective study. Within 2 hours each patient was examined by 5 different physicians using 2 of 3 different, randomly allocated ultrasound machines. Each investigator determined the hilar and
Introduction

Sonography plays an important role in follow-up examinations after kidney transplantation to monitor transplant function and to detect early signs of transplant rejection [1–3], renal venous thrombosis [4, 5], stenosis of the renal artery or arterial anastomosis [6] and delayed graft function [7]. Determining the renal resistive index (RRI) of renal parenchyma, i.e., arcuate or interlobar, arteries and renal hilar arteries is the mainstay of the examination. The arterial RRI is defined as: (peak systolic velocity – end diastolic velocity)/peak systolic velocity and a value around 0.6 is considered normal with an upper boundary in renal allografts of 0.8 [8] or 0.9 [9]. Additionally, sudden changes in the RRI may herald an impending problem.

A plethora of studies are available that have investigated the RRI as a marker of renal function comparing it with serum creatinine levels or histologic findings and as a predictor of postoperative rejection [4, 5], renal venous thrombosis [6–15], delayed graft function [7]. Discrepancies between study results sparked a controversy about the validity of the RRI as an indicator of purely renal organ damage. More and more systemic cardiovascular factors have since been identified that have an impact on the RRI, which is gradually leading to a more reasonable interpretation of the RRI, thereby helping to restore its applicability [10, 16, 17].

In this ongoing debate, comparatively little data is available that affirms the reproducibility of the measurement itself, i.e., the purely technical aspects of the examination. One cannot help but notice that there seems to be a certain degree of arbitrariness when measuring the RRI resulting from the quality of the velocity-time spectral recordings obtained, the subjectivity in defining peak systolic and end diastolic velocities, and the fact that it is unlikely that the same renal arcuate or interlobar artery is being measured in serial examinations. While some patient- or setting-related factors cannot be altered, there are some well-known errors the examiner can avoid to improve the quality of the signal waveform. These errors comprise mistaking background noise for real flow when gain is set too high, removing actually present low-frequency Doppler signals with an incorrect wall filter setting, creating aliasing artifacts or a small waveform when an inadequate pulse repetition frequency is used, and obtaining a mixture of signals from different vessels in close proximity to each other if the sampling volume is too large [18–20].

At our institution, follow-up examinations after kidney transplantation are conducted according to a rather rigid scheme with daily examinations during the first 7 to 10 days after transplantation, followed by a check-up on a weekly basis until settling into a routine check-up once a year beginning with the second year after kidney transplantation. Any abnormalities, signs of kidney failure or transplant rejection warrant additional examinations. It is evident that this procedure requires a lot of resources on the part of the health care provider and is time-consuming for the patient. It also poses a risk of infection, especially in the posttransplant period when the patient’s wound are not yet fully healed. It is also clear that these examinations are unlikely to be performed by the same physician or with the same ultrasound machine as institutions performing kidney transplants are of a certain minimum size. Reproducibility of a measurement is a prerequisite for its validity as a diagnostic tool. If this is not the case, such resource-intensive and time-consuming follow-up schemes might need to be reconsidered.

The aim of this study therefore was to assess the reproducibility of RRI measurements from a technical perspective by evaluating its consistency for different ultrasound machines, examiners, and patient characteristics that might possibly hamper the ultrasound examination.

Materials and Methods

Patients

The study was approved by the local ethics committee. Over a period of 4 months, 22 consecutive patients who were scheduled for a...
follow-up examination of their kidney transplant at our department and were willing to participate in the study were prospectively included. Written informed consent from all patients was obtained prior to participation in the study. Patients were either inpatients who had just received a kidney transplant within the previous couple of weeks or outpatients with a kidney transplant of any age. Each patient was examined by the physician on sonography duty and immediately afterwards by 4 additional observers, resulting in 5 examinations per patient and 110 examinations altogether. Creatinine level and glomerular filtration rate (GFR) were retrieved from the laboratory test of the same day. These patient characteristics are summarized in Table 1.

Observers

19 physicians from the department of radiology with experience ranging from 6 months of sonography as a radiologist in training to board-certified specialized radiologist with a particular interest in sonography participated in this study. Apart from being a licensed physician, additional requirements to take part in this study as an observer were to have at least 6 months of experience in sonography and to be able to independently perform clinical sonographic examinations without direct supervision. Observers were categorized into two groups by experience level: very experienced (physicians working in the radiology department for at least 4 years) and moderately experienced (all others). All observers were blinded to the findings of previous examinations, the patient’s current clinical status, and the RRI measurements of other examiners.

Observers participated in the study when available at the time an eligible patient showed up for the examination. Care was taken that out of the 4 observers who examined one patient at a given time 2 were very experienced and 2 were moderately experienced while the routine examination of the patient was performed by the physician on sonography duty on the given day, which tended to be a physician with moderate experience in sonography.

Ultrasound machines

Three different ultrasound machines were used in this study: Ultrasound machine 1: GE, LOGIQ E9 with convex probe transducer C1–6 (1.5–6 MHz); ultrasound machine 2: Siemens, ACUSON X700 (Model 10 658 844) with convex probe transducer Acuson (1–4.5 MHz); ultrasound machine 3: Toshiba, Apio XG (SSA-790A) with convex probe transducer PVT-375BT (3.5 MHz).

Each patient was examined by 5 different physicians immediately after each other, with 2 different ultrasound machines being used for these 5 examinations. Assignment to the ultrasound machine was done randomly, always using the next machine available.

Sonography with RRI measurement

The convex probe transducer and the standard preset for abdominal examinations of the given ultrasound machine were used. After B-mode sonography to localize the kidney allograft, the color Doppler function was used to detect hilar, interlobar and arcuate arteries, which were then analyzed in the pulsed-wave Doppler mode to generate a velocity-time curve. The pulse repetition frequency (PFR) was initially set to 1.5 kHz and the sample volume to 2 mm. Each examiner adjusted the parameters as he or she saw fit and would have done in any routine examination. The peak systolic and end diastolic velocities were automatically determined and manually corrected as deemed necessary by the examiner. The RRI was calculated according to the formula: (peak systolic velocity – end diastolic velocity)/peak systolic velocity. The three RRIs derived from the arcuate or interlobar arteries were averaged and defined as the parenchymal RRI.

The shortest distance between the body surface and the renal allograft was measured in B-mode sonography to obtain a value for the depth of the allograft within the recipient’s body, a parameter included in further analysis because of its possible impact on reproducibility.

Statistical analysis

The reproducibility and reproducibility limit of the RRI were determined as described by Strouthidis et al. [21], and Cronbach’s alpha and the intraclass correlation coefficient (ICC) were analyzed. Deviation of the RRI from the mean RRI over the 5 measurements was used as an indicator of reproducibility. The impact of different ultrasound machines, examiner’s experience level, renal function impairment (GFR < 45 ml/min), and inpatient versus outpatient status was tested with the Kruskal-Wallis test. The bivariate linear correlation of the minimal transplant distance from the body surface with the variance of the parenchymal RRI was performed as well as correlation analysis of various factors with the average RRI. Differences between the two groups with respect to patient characteristics were assessed with the Mann-Whitney U (MWU) test. All statistical analysis was performed using SPSS Statistics 23 (IBM, Armonk, NY, USA) and GraphPad Prism (Graph Pad Soft-
ware, Inc.). A P-value of < 0.05 was considered statistically significant. Bland-Altman plots were generated with GraphPad Prism (Graph Pad Software, Inc.).

Results

Bivariate correlation analysis showed at most a minimal tendency of the hilar and parenchymal RRI to correlate positively with serum creatinine levels with a Pearson correlation coefficient (PCC) of 0.412 and p = 0.057 for the hilar RRI and a PCC of 0.346 with p = 0.115 for the parenchymal RRI and negatively with GFR with a PCC of -0.343 and p = 0.118 for the hilar RRI and a PCC -0.345 with p = 0.116 for the parenchymal RRI. None of these correlations reached statistical significance. There was also no correlation of RRIs with patient age (p = 0.331 for hilar RRI and p = 0.499 for parenchymal RRI).

Despite an uneven gender distribution with a greater proportion of men, no statistically significant difference of hilar or parenchymal RRIs was found between men and women (Mann-Whitney U-test p = 0.538).

As there is no gold standard for the RRI, we used the deviation of the RRI determined in an examination from the mean RRI over the 5 measurements of the same patient performed within 2 hours as a surrogate for reproducibility. The Bland-Altman plots in ▶ Fig. 1a, b illustrate the agreement between the mean of the RRI measurements and the individual RRI measurements.

In our study the reproducibility (SR) of the RRI measurements was found to be 0.045 with a reproducibility limit (R) of 0.124. Reliability analysis of RRI measurements yielded an intraclass correlation coefficient (ICC, one-way random) of 0.848 and a 95% confidence interval of 0.718–0.929 for the parenchymal RRI and of 0.865 with a 95% confidence interval of 0.751–0.937 for the hilar RRI (▶ Table 2).

A statistically significant difference in the degree of deviation of the parenchymal RRI measurements from the mean RRI and the type of ultrasound machine used was found with the Kruskal-Wallis test (p = 0.003) while the examiners’ experience levels, kidney function impairment defined as a glomerular filtration rate (GFR) of <45 ml/min, or inpatient versus outpatient status did not show a statistically significant impact (▶ Table 3, ▶ Fig. 2).

The variance of the parenchymal RRI correlated significantly with minimal renal allograft distance from the body surface with a Pearson’s correlation coefficient of 0.649 and a significance of p = 0.001 (▶ Fig. 3).

Discussion

Bivariate correlation analysis showed at most a minimal tendency of hilar and parenchymal RRI to correlate positively with serum creatinine levels and negatively with GFR without reaching statistical significance. This was somewhat expected considering that we did not include any of the already established systemic factors with a known impact on RRI in our study. The influence of these systemic factors on the RRI is nicely demonstrated elsewhere [8]. However, this study primarily aimed at testing the reproducibility of the RRI measurement as such and not its correlation with renal function tests.

▶ Fig. 1 a Bland-Altman plot of parenchymal renal resistive index (RRI) measurements in comparison to the mean parenchymal RRI over 5 observers. b Bland-Altman plot of hilar renal resistive index (RRI) measurements in comparison to the mean hilar RRI over 5 observers.

▶ Abb. 1 a Bland-Altman-Diagramm der parenchymalen Widerstandsindexmessungen verglichen mit dem Mittelwert der Messungen der 5 Untersucher. b Bland-Altman-Diagramm der hilären Widerstandsindexmessungen verglichen mit dem Mittelwert der Messungen der 5 Untersucher.

▶ Table 2 Reliability analysis for 5 consecutive examinations. RRI: renal resistive index; ICC: intraclass correlation coefficient; CI, confidence interval.

<table>
<thead>
<tr>
<th></th>
<th>Cronbach’s alpha</th>
<th>ICC one-way random</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parenchymal RRI</td>
<td>0.852</td>
<td>0.848</td>
<td>0.718–0.929</td>
</tr>
<tr>
<td>Hilar RRI</td>
<td>0.868</td>
<td>0.865</td>
<td>0.751–0.937</td>
</tr>
</tbody>
</table>

Theilig DC et al. The Renal Resistive... Fortschr Röntgenstr
Reliability analysis of RRI measurements yielded rather good reproducibility in our study with an ICC of 0.848 and a 95% confidence interval of 0.718–0.929 for the parenchymal RRI and an ICC of 0.865 with a 95% confidence interval of 0.751–0.937 for the hilar RRI. This agrees with the sparsely available previous studies [22, 23], which, however, did not include as many different examiners or different types of ultrasound machines or reflect typical clinical settings.

The reproducibility (SR) of the RRI measurements in our study was found to be 0.045 with a reproducibility limit (R) of 0.124 which is a useful orientation value for the interpretation of changes in RRI in a routine clinical setting. In other words, according to our study, changes in RRI measurements need to be greater than 0.124 apart to assume that the measurements truly differ from each other, as changes in RRI less than 0.124 are with a likelihood of 95% within the range of measurement variation.

Experienced sonography readers might object that our study simply reflects the aptitude of individual examiners and possibly their degree of familiarity with the different ultrasound machines and that they themselves would achieve a better reproducibility. This might indeed be true. However, the aim of our study was to test reproducibility and to obtain the reproducibility limit of the technique in a realistic setting, i.e., whether it is robust enough to be relied upon in the scenarios typically encountered, which is most likely different physicians with variable ultrasound skills using different types of ultrasound machines for follow-up examinations and not an expert with a personal ultrasound machine and individualized presets. Such an expert is typically not around in the middle of the night when the first check-up after kidney transplantation is due. With 19 different physicians having participated in this study with an experience ranging from 6 months of sonography as a radiologist in training to board-certified specialized radiologist with a particular interest in sonography, we are confident that this study reflects a representative cohort of examiners. Our study, therefore, underlines the validity of RRI measurements in a typical clinical setting.

Reliability analysis of RRI measurements yielded rather good reproducibility in our study with an ICC of 0.848 and a 95% confidence interval of 0.718–0.929 for the parenchymal RRI and an ICC of 0.865 with a 95% confidence interval of 0.751–0.937 for the hilar RRI. This agrees with the sparsely available previous studies [22, 23], which, however, did not include as many different examiners or different types of ultrasound machines or reflect typical clinical settings.

The reproducibility (SR) of the RRI measurements in our study was found to be 0.045 with a reproducibility limit (R) of 0.124 which is a useful orientation value for the interpretation of changes in RRI in a routine clinical setting. In other words, according to our study, changes in RRI measurements need to be greater than 0.124 apart to assume that the measurements truly differ from each other, as changes in RRI less than 0.124 are with a likelihood of 95% within the range of measurement variation.

Experienced sonography readers might object that our study simply reflects the aptitude of individual examiners and possibly their degree of familiarity with the different ultrasound machines and that they themselves would achieve a better reproducibility. This might indeed be true. However, the aim of our study was to test reproducibility and to obtain the reproducibility limit of the technique in a realistic setting, i.e., whether it is robust enough to be relied upon in the scenarios typically encountered, which is most likely different physicians with variable ultrasound skills using different types of ultrasound machines for follow-up examinations and not an expert with a personal ultrasound machine and individualized presets. Such an expert is typically not around in the middle of the night when the first check-up after kidney transplantation is due. With 19 different physicians having participated in this study with an experience ranging from 6 months of sonography as a radiologist in training to board-certified specialized radiologist with a particular interest in sonography, we are confident that this study reflects a representative cohort of examiners. Our study, therefore, underlines the validity of RRI measurements in a typical clinical setting.

The examiner’s degree of experience did not have an impact on the reproducibility of the RRI measurement in our study, which is in contradiction with the findings of Mikkonen et al. [24]. This could be explained by the fact that we did not define a gold standard of RRI measurement as there is none, but rather compared measurements with each other and the mean, assuming that a larger variance between the measurements signifies a lesser degree of reproducibility and vice versa. The impact of a potentially
more precise measurement by a more experienced examiner might therefore well have been negated by the less precise measurements of the less experienced examiners.

Factors that we did find to have a statistically significant impact on the reproducibility of the RRI were the type of ultrasound machine and the distance of the kidney transplant from the body surface.

With regard to the differences between the types of ultrasound machines used, it has to be said that, while statistically significant, the differences were rather small with the deviation of the parenchymal RRI measurement from the mean RRI differing only by 0.02 between the types of ultrasound machines. A possible explanation for these discrepancies might be differences in the display of the velocity-time spectral recordings due to the varying algorithms used and graphical aspects that might have led to a less accurate manual cursor adjustment after automatic determination of the peak systolic and end diastolic velocities with some machines. Furthermore, the ultrasound machines used were of different generations, the most accurate results where indeed obtained with the most modern machine, the GE LOGIQ E9.

Notably but not surprisingly, the reproducibility of the RRI is hampered by a deep location of the kidney transplant within the body, and this technical limitation should, therefore, prompt the clinician to refrain from relying on RRI measurements in such a setting.

The study has some limitations, namely the rather small number of patients investigated and the above-mentioned fact that is the case that may well have been negated by the less precise measurement by a more experienced examiner.

Particularly patients in intensive care units are often looked after by various physicians and, depending on the hospital structure, even different specialties, i.e., surgeons, nephrologists, urologists, and radiologists. Our results emphasize that providing an ultrasound machine in intensive care units that is used by all physicians involved in patient care is important in order to improve the comparability of measurements and to minimize technical factors potentially obscuring RRI assessment.

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgements

The authors thank Bettina Herwig for language editing.

References


Theilig DC et al. The Renal Resistive... Fortschr Röntgenstr

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.