Whole-body magnetic resonance imaging in the large population-based German National Cohort study: predictive capability of automated image quality assessment for protocol repetitions

Schuppert, C., von Krüchten, R., Hirsch, J.G., Rospleszcz, S., Hoinkiss, D.C., Selder, S., Köhn, A., von Stackelberg, O., Peters, A., Völzke, H., Kröncke, T., Niendorf, T., Forsting, M., Hosten, N., Hendel, T., Pischon, T., Jöckel, K.H., Kaaks, R., Bamberg, F., Kauczor, H.U., Günther, M., Schlett, C.L.

The following statement is required by the publisher: "This is a non-final version of an article published in final form in: Investigative Radiology 57(7):p 478-487, July 2022.
DOI: 10.1097/RLI.0000000000000861"

This is the final version of the accepted manuscript. The original article has been published in final edited form in:

Investigative Radiology
2022 JUL 01 ; 57(7): 478-487
2022 FEB 17 (originally published online)
DOI: 10.1097/RLI.0000000000000861

Publisher: Wolters Kluwer Health | Lippincott Williams & Wilkins

Whole-Body Magnetic Resonance Imaging in the Large Population-Based German National Cohort Study: Predictive Capability of Automated Image Quality Assessment for Protocol Repetitions

Christopher Schuppert, MD*[1]; Ricarda von Krüchten, MD*[2]; Jochen G. Hirsch, PhD[3]; Susanne Rospleszcz, PhD[4,5,6]; Daniel C. Hoinkiss, PhD[3]; Sonja Selder, PhD[7]; Alexander Köhn, MSc[3]; Oyunbileg von Stackelberg, PhD[1]; Annette Peters, PhD[4,5,6]; Henry Völzke, MD[8,9]; Thomas Kröncke, MD MBA[10]; Thoralf Niendorf, PhD[11]; Michael Forsting, MD[12]; Norbert Hosten, MD[13]; Thomas Hendel, PhD[5]; Tobias Pischon, MD MPH[14]; Karl-Heinz Jöckel, PhD[15]; Rudolf Kaaks, PhD MSc[16]; Fabian Bamberg, MD MPH[2]; Hans-Ulrich Kauczor, MD[1]; Matthias Günther, PhD[3]; Christopher L. Schlett, MD MPH[2];

for The German National Cohort MRI Study Investigators

[1]Department of Diagnostic and Interventional Radiology, Heidelberg University Hospital, Heidelberg, Germany

[2]Department of Diagnostic and Interventional Radiology, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Germany

[3]Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany

[4]Institute of Medical Information Processing, Biometry and Epidemiology, Ludwig Maximilian University, Munich, Germany

[5]Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[6]German Centre for Cardiovascular Research, Partner Site Munich Heart Alliance, Munich, Germany

[7]Department of Radiology, Ludwig Maximilian University Hospital, Munich, Germany

[8]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany

[9]German Centre for Cardiovascular Research, Partner Site Greifswald, University Medicine Greifswald, Greifswald, Germany

[10]Department of Diagnostic and Interventional Radiology, University Hospital Augsburg, University of Augsburg, Germany

[11]Berlin Ultrahigh Field Facility, Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany

[12]Department of Diagnostic and Interventional Radiology and Neuroradiology, University Hospital Essen, University Duisburg-Essen, Essen, Germany

[13]Department of Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany

[14]Molecular Epidemiology Research Group, Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany

[15]Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University Duisburg-Essen, Essen, Germany

[16]Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany

[*]equal contribution

Corresponding Author

Christopher L. Schlett, MD MPH

Department of Diagnostic and Interventional Radiology

Medical Center – University of Freiburg

Hugstetter Str. 55, 79106 Freiburg, Germany

christopher.schlett@uniklinik-freiburg.de

ABSTRACT

Background: Reproducible image quality is of high relevance for large cohort studies and can be challenging for magnetic resonance imaging (MRI). Automated image quality assessment may contribute to conducting radiologic studies effectively.

Purpose: The aims of this study were to assess protocol repetition frequency in population-based whole-body MRI along with its effect on examination time and to examine the applicability of automated image quality assessment for predicting decision-making regarding repeated acquisitions.

Materials and Methods: All participants enrolled in the prospective, multicenter German National Cohort (NAKO) study who underwent whole-body MRI at 1 of 5 sites from 2014 to 2016 were included in this analysis (n = 11,347). A standardized examination program of 12 protocols was employed. Acquisitions were carried out by certified radiologic technologists, who were authorized to repeat protocols based on their visual perception of image quality. Eleven image quality parameters were derived fully automatically from the acquired images, and their discrimination ability regarding baseline acquisitions and repetitions was tested.

Results: At least 1 protocol was repeated in 12% (n = 1359) of participants, and more than 1 protocol in 1.6% (n = 181). The repetition frequency differed across protocols ($P < 0.001$), imaging sites ($P < 0.001$), and over the study period ($P < 0.001$). The mean total scan time was 62.6 minutes in participants without and 67.4 minutes in participants with protocol repetitions (mean difference, 4.8 minutes; 95% confidence interval, 4.5-5.2 minutes). Ten of the automatically derived image quality parameters were individually retrospectively predictive for the repetition of particular protocols; for instance, “signal-to-noise ratio” alone provided an AUC of 0.65 ($P < 0.001$) for repetition of the Cardio Cine SSFP SAX protocol. Combinations generally improved prediction ability, as exemplified by “image sharpness” plus “foreground ratio” yielding an area under the curve of 0.89 ($P < 0.001$) for repetition of the Neuro T1w 3D MPRAGE protocol, versus 0.85 ($P < 0.001$) and 0.68 ($P < 0.001$) as individual parameters.

Conclusion: Magnetic resonance imaging protocol repetitions were necessary in approximately 12% of scans even in the highly standardized setting of a large cohort study. Automated image quality assessment shows predictive value for the technologists' decision to perform protocol repetitions, and has the potential to improve imaging efficiency.

Keywords

Population imaging; Large cohort study; Magnetic resonance imaging; Body imaging; Image quality; Technical image analysis; Workflow optimization; Radiologic technologist; Epidemiology

Abbreviations

AUC	Area under the curve
CI	Confidence Interval
DNN	Deep neural network
FOV	Field of view
MRI	Magnetic resonance imaging
NAKO	German National Cohort
UQI	Universal quality index
RT	Radiologic technologist
SNR	Signal-to-noise ratio

INTRODUCTION

Whole-body magnetic resonance imaging (MRI) has become a primary imaging technique in population-based cohort studies due to its excellent spatial resolution and soft tissue contrast, capacity for standardization, and lack of ionizing radiation. Several population-based cohort studies used this technique, such as the Multi-Ethnic Study of Arteriosclerosis [1], the Framingham Heart Study [2], and the Study of Health in Pomerania [3], among others. In addition, 2 large ongoing population-based studies, namely the UK Biobank (100,000 participants) [4] and the German National Cohort (NAKO; 30,000 participants) [5], rely on this modality. Their comprehensive databases that combine detailed imaging and nonimaging phenotyping can provide valuable information about general health, and thus are a valuable source for assessing potential risk markers, as well as identifying radiomic features of subclinical disease states and personalized medicine [6], or ascertaining the prevalence of incidental findings and outcomes [7].

Standardized and reproducible image acquisition is indispensable in such large cohort studies to warrant a consistent basis for further postprocessing, including automated segmentation tasks and data extraction [8]. The examination programs used in whole-body MRI are usually complex and tedious for both radiologic technologists (RTs) and participants, who may be immobilized for the duration of examination, depending on the study [5]. The RTs' workflow generally involves reviewing all images immediately postacquisition to assess quality and to repeat a series if deemed necessary. Common reasons for repetition are image blurring (bulk or physiological motion artifacts), distortions due to susceptibility artifacts, low signal and/or high noise, incorrect anatomical coverage, or protocol-specific abnormalities such as fat-water swapping due to faulty shimming or center frequency settings. This decision-making process is time-consuming and relies on subjective visual perception and professional experience. Moreover, prolonged scan times can result in additional discomfort and increase the likelihood of participant dropouts. In the context of large cohort studies, this process then becomes both inefficient and costly. Developments

leading to automation in workflow and RTs' decision-making are therefore being undertaken and could result in improved resource allocation and image quality, as well as less discomfort for participants and patients [9, 10].

The fundamental role of RTs' perception of image quality and its influence on protocol repetition in a large cohort study, as a preceding factor to all further analyses, has not previously been established in the literature. Our study, therefore, aimed to assess protocol repetition frequency as well as the RTs' underlying decision-making process, to determine its correlation with quantitative image quality parameters, and to identify quantitative parameters that are predictive for protocol repetition, with a perspective to objectify and potentially automate this time-consuming task.

MATERIALS AND METHODS

Study Design and Population

Our project was designed as a predefined ad hoc analysis on data from the MRI study of the NAKO. The NAKO is an ongoing, prospective, interdisciplinary, multicenter, population-based cohort study undertaken by a network of over 25 German institutions and spans 18 study centers across Germany. Its main goal is to investigate risk factors for the development of common chronic diseases such as cancer, diabetes, cardiovascular, neurodegenerative/psychiatric, respiratory, and infectious diseases [11]. For at least 25 years, more than 200,000 participants of the general population between the ages of 20 and 69 years will be examined in a baseline and follow-up studies. Besides interviews, questionnaires, physical examinations, as well as the collection of biological samples, a subgroup of 30,000 participants received a baseline whole-body MRI at 1 of 5 dedicated imaging centers across Germany between May 2014 and April 2019 [5]. For our study, we included all participants enrolled in the MRI substudy until December 31, 2016, which resembles all currently available data. Participants were excluded from the analysis if they were early dropouts (the examination was interrupted before the first protocol was fully acquired) or if the participants withdrew consent. The scientific advisory board and the ethics advisory board of the NAKO approved this study.

Image Acquisition

MRI was performed using identical 3 T whole-body scanners (MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany) installed at the 5 imaging centers, running an identical software version. A high field strength system was chosen for its improved signal-to-noise ratio and enhanced spatial resolution, despite higher costs and a higher likelihood of artifacts. The examination program comprised a set of 12 non-contrast agent-enhanced series, covering 4 anatomical domains, to form a whole-body examination **(Figure 1)**. Aimed at the evaluation of neurological, cardiovascular, thoracoabdominal, and musculoskeletal

pathologies, as well as subclinical disease burden and physiological variants, the domains were each allocated 15 minutes of scan time, with a total scan time of approximately 60 minutes. A detailed account of the rationale, design and technical background of the MRI substudy has been given previously [5].

All acquisitions were carried out by RTs who received training and certification specifically for the MRI substudy and had to be recertified annually. Participants were briefed in detail on the scanning procedure and cautioned to minimize movement as well as follow breathing instructions. RTs were instructed to repeat measurements if anatomical coverage was not adhering to an internally defined standard, if severe artifacts were present, or if they considered image quality to be otherwise unsatisfactory.

Automated Image Quality Assessment

All acquired images were transferred “on-the-fly”, parallel to the ongoing examination, from the imaging centers to a central storage facility, dubbed the “imaging core”, using a virtual private network and the standard DICOM (Digital Imaging and Communications in Medicine) format. Besides general data management, the imaging core provided basic automated quality assurance, including checks for data completeness, conformity to the predefined protocol parameters, and data uniqueness. A set of 11 image-based quality parameters was then calculated automatically from the acquired images (calculation time slightly below 60 minutes for 1 complete examination). It included a proprietary universal quality index (UQI), sharpness, signal-to-noise ratio (SNR) and specific SNR, structured noise maximum and average, N/2 ghosting level maximum and average, drift, variation over time, and foreground ratio.

The UQI served as a first indication of image quality without giving specific characteristics. It uses the original image, a noise-filtered, and an edge-filtered version to determine a score that increases with image noise and decreases with image blur. Sharpness was evaluated using an entropy focus criterion, based on previously published

141 work [12]. Further image quality assessment required the computation of binary foreground
142 and background masks; these were generated using established thresholding methods [13-
143 15], although robust mask generation was not possible with all protocols. SNR was then
144 calculated based on existing methodology [16]. Because its calculation considers the whole
145 image background, the result may be affected by artifacts there, for example, caused by
146 subject motion. For this reason, a specific SNR was additionally calculated using predefined
147 regions of interest to only consider real image noise.
148 For the detection of structured noise and nyquist-ghosting artifacts, iterative
149 line/column correlation algorithms were used. The first examined the correlation between
150 neighbored lines/columns to detect structure in the image background, which could be
151 caused, for example, by streaking artifacts. The latter separated these lines/columns by half
152 the size of the field of view (FOV) to detect nyquist-ghosting. We assessed the results by
153 averaging over all correlations for general image quality, as well as selecting the maximum
154 value for finding strong local artifacts.
155 The position of the binary foreground mask was used to evaluate the positioning of
156 the subject. The ratio between foreground and background mask areas then represents the
157 ratio of subject size to FOV choice. Lastly, low-frequency signal drifts are often observed in
158 functional MRI with demanding imaging sequences [17]. A dedicated test module quantified
159 this intensity drift for the Resting State EPI BOLD protocol by calculating the time-wise
160 changes in mean signal intensity in the image foreground, as well as the signal variance
161 over the time series.
162 For repeated measurements, RTs were required to select one acquisition that would
163 primarily be used for reading and research, while provided with the information above
164 through a Web-based thin client. An additional quality check was performed manually for
165 each protocol by NAKO investigators (radiology residents and board-certified radiologists)
166 while reading for incidental findings. Using a traffic light rating system, protocols were
167 categorized based on completeness of anatomical coverage and severity of artifacts. Those
168 labeled “red” were excluded from further analyses.

Statistical Analysis

Data are presented as mean with standard deviation for continuous variables and counts and percentages for categorical variables. Differences in image quality parameters between initial protocol acquisitions with and without subsequent repetition were compared by Student *t* test. Discrimination ability was assessed by receiver operating characteristic curves and corresponding area under the curve (AUC) for single image quality parameters as well as for explorative combinations of multiple parameters. A *P* value of < 0.05 was considered to denote statistical significance. Statistical analysis was performed using SAS (version 9.4; SAS Institute Inc., Cary, NC, USA) and the R programming environment (version 3.6.3; R Foundation, Vienna, Austria).

RESULTS

A total of 11,347 participants were examined in the specified time frame. A majority of 10,960 (96.6%) completed the full MRI examination. Recruitment across the 5 imaging centers varied within a range of 1319 to 3093 participants per site. Most participants were examined on weekdays; a minority of 350 participants (3.1%) underwent imaging on Saturdays. Nearly equal distribution of participants was observed for the morning (07:00–11:59 AM) and afternoon (12:00–4:59 PM) scanning periods, accounting for 47.0% and 45.8% of the study sample, while the evening period (5:00 PM and later) comprised a minority of 7.2%. Overall mean scan time was 63.1min and 63.7min, if restricted to participants with complete examinations.

MRI Protocol Repetition—A Per Participant Analysis

In 1359 participants (12.0%), a total of 1558 different protocols were repeated at least once based on the real-time decision of the RT. Repetitions were limited to only 1 protocol for the majority of 1178 participants (10.4%), whereas 2 or more different protocols were repeated in 181 participants (1.6%): 2 in 165, 3 in 14, and 4 protocols in 2 participants. The mean total scan time increased by 4.8 minutes (95% confidence interval (CI), 4.5–5.2 minutes; SD, ±5.9 minutes) from 62.6 to 67.4 minutes in participants with 1 or more protocol repetitions.

Repetition frequency in participants (from here on defined as the percentage of participants with at least 1 protocol repetition) fluctuated over the enrollment period with a range of 4.5% to 15.3% in 3-month intervals ($P < 0.001$), and simultaneously varied across sites (range, 2.3–28.1% sitewide for the same periods, $P < 0.001$) (**Figure 2**). A higher repetition frequency was observed on Saturdays (16.3%) versus weekdays (11.8%, $P = 0.01$). Differences in repetition frequency between daytimes were not significant (11.6% vs 12.1% vs 14.2% for morning [7:00 to 11:59 AM], afternoon [12:00 to 4:59 PM], and evening [5:00 to 9:30 PM], $P = 0.11$). In a multivariable analysis, only the study site and the enrollment period (time since the first examination) remained predictive for differences in

repetition frequency, and notably, odds for protocol repetition increased by 1.25-fold (95% CI, 1.13–1.39) per year of enrollment.

MR Protocol Repetitions—A Per Protocol Analysis

Of all initially acquired protocols (n = 134,239), 1.2% (n = 1558) were repeated at least once. The frequency ranged between 0.1% and 3.6% per protocol; the most frequently repeated protocols were Cardio Cine SSFP LAX nK (3.6% of initial acquisitions were repeated), Neuro 2D FLAIR axial (2.9%), and Body Multi-echo 3D VIBE (1.9%), while the least frequently repeated protocols were Cardio MOLLI SAX (0.1%) and Neuro Resting State (0.1%) (**Figure 3**). Merely 0.03% of initial protocol acquisitions were repeated more than once, resulting in a total of 1606 repetition scans.

Automatically Derived Image Quality Parameters Predicting MR Protocol Repetition

Ten of the automatically derived image quality parameters, if considered individually, exhibited statistically significant differences between initial acquisitions with and without subsequent repetition **(Table 1** and **Supplemental Table 1**, exemplary images and distribution plots; **Figure 4)**. The 2 parameters demonstrating this discriminative behavior for the largest number of protocols were image sharpness and SNR: in 9 of 14 protocols for which it was measured, sharpness differed significantly between initial and repeated protocols ($P < 0.001$ to $P = 0.049$), and SNR differed significantly in 6 of the 12 protocols for which it was calculated ($P < 0.001$ to $P = 0.003$). For 3 protocols, statistically significant differences between the 2 groups were not observed for any image quality parameter: MSK-Spine T2w FSE sagittal, Body T2w HASTE axial, and Cardio MOLLI SAX.

Classification performance of the image parameters in terms of AUC was especially high for the neurological protocols, with several parameters exhibiting areas greater than 0.75 when comparing initial acquisitions with and without subsequent repetition **(Table 1)**. Combinations of image quality parameters generally improved discriminative ability over single parameters **(Supplemental Table 2)**, although performance remained notably poor

for the musculoskeletal protocols MSK-Hip PDw FS 3D SPACE and MSK-Spine T2w FSE sagittal **(Figure 5)**.

Setup Changes for Protocol Repetition and their Effect on the Discriminative Power of Automatically Derived Image Quality Parameters

Approximately half of the 1606 protocol repetitions (49.8%) were carried out with manual adjustments to the technical setup, namely, radiofrequency (RF) coil configuration (variations in RF coils and in selection of receive RF coil elements), FOV size, slice position (FOV shifted along the *x/y/z* axis of the participant), or slice orientation (FOV rotated or angled differently). A closer examination revealed that about one third of repetitions (32.4%) involved changes to a single attribute, predominantly slice position (31.9%). If 2 attributes were altered (14.1%), these concerned primarily slice position and slice orientation in combination (9.5%). Field of view adjustments were always accompanied by additional changes **(Table 2**, detailed breakdown between protocols: **Supplemental Table 3)**. Comparing these adjustments with the examination guidelines, they were generally made due to an incorrect or at least inferior initial setup.

The other half of protocol repetitions (50.2%) was performed without any manual adjustments to the technical setup, leading to the assumption that the RT operated solely on the grounds of subjectively low image quality, and did not attribute their subpar visual impression to any changeable technical factor.

Across all protocols, there was a clear negative correlation between the proportion of manual setup changes and maximum classification ability as measured by AUC (**Figure 6**). The correlation coefficient was $r = –0.30$ ($r^2 = 0.09$), and decreased to $r = –0.67$ (95% CI, –0.90 to –0.13; $r^2 = 0.46$) after removal of the outlying AUC for the functional Neuro Resting State protocol, thereby demonstrating that the fewer manual setup changes were performed, the better image quality parameters were able to discriminate between baseline acquisitions and repetitions.

DISCUSSION

In our sample of the MRI substudy of the population-based NAKO study, 11,347 participants underwent whole-body MRI. Of these, 10,960 (96.6%) completed the entire MRI sequence with a mean scan time of 63.1 minutes. In 1359 participants (12.0%), the RT decided to repeat at least 1 protocol, due to incorrect technical setup parameters or based on the subjective impression of insufficient image quality. Automatically derived image quality parameters were able to discriminate between the baseline acquisitions and repetitions.

In general, the decision for protocol repetition will substantially depend on the RTs' level of experience and expertise, especially if subjectively low image quality occurs despite an objectively error-free technical setup. In a clinical setting, achieving optimal examination quality by well-selected protocol repetition is instrumental to patients receiving precise and correct diagnoses, and may influence therapeutic outcomes. Moreover, protocol repetition is a time-costly process, verified by our results, where the mean scan time increased by 4.8 minutes or 7.7% for examinations with protocol repetitions, which amounted to an additional 94.2 hours of scan time (or 11.8 days, considering an 8-hour workday). When considering even larger cohort studies or routine clinical practice, both the time losses and the cost burden become considerable. Equally important is the effect of increased scan time on study participants and patients, who may already be in physical or psychological distress, and could benefit from shorter examinations.

Interestingly, in a breakdown of 3-monthly intervals, the repetition frequency fluctuated over the enrollment period, ranging from a low of 4.5% to a high of 15.3%, and simultaneously varied across sites from 2.3% to 28.1%. These observations remained significant in a multivariate analysis and yielded 1.25-fold per year increased odds for protocol repetition. This may be explained by a continuous learning process, in which the RT increased their ability to identify suboptimal image quality over time, and therefore more likely performed protocol repetitions as they gained professional experience within the NAKO study. A close site monitoring by the NAKO administration for quality assurance,

personalized feedback to RTs, and yearly training refreshers that highlighted common quality issues may have been contributing factors as well. We did not, however, identify specific factors for the variations between imaging sites.

The most frequently repeated protocol was Cardio Cine SSFP LAX nK, known to be susceptible to off-resonance or banding artifacts as well as incorrect slice positioning, and possibly additionally error-prone in this particular study due to being among the last protocols in a time-intensive 60-minute scanning program. On the other hand, the least frequently repeated protocols were Cardio MOLLI SAX and Neuro Resting State EPI BOLD; both of which are not as intuitively assessable as the rest of the acquired protocols, which possibly contributed to their low rate of repetition. We were, however, unable to support this claim with the study data, as the RTs' reasoning regarding protocol repetition was not explicitly documented.

We identified differences in the predictive ability of the automatically derived image quality parameters regarding protocol repetition and found parameter combinations to be more effective than single parameters. Their predictive ability was particularly good for protocols in the neurological domain—this could be the result of parameters being explicitly “pure” in a static body region not prone to motion (breathing) artifacts, although less obvious factors may have contributed as well. It could also imply an above-average proficiency of the RTs in neuroradiological imaging over the remaining domains.

Considering previous study findings as well as our own observations of the workflow challenges faced by technologists, the advantages of automated image quality assessment become more evident. In a substudy of the UK Biobank cohort, 100,000 participants underwent cardiovascular MRI without the implementation of automated image quality control. This was tedious and time-consuming for RTs, who had to undergo multiple training sessions to ensure consistent quality assessment [18]. Five years later, in the same cohort, an automated image quality control using machine learning methods was tested on the first 10,000 brain imaging datasets and evaluated against a validation set of manually assessed examinations in 5816 participants [19]. The performance of the algorithm was satisfactory

and subsequently reduced the need for manual checking. Particularly algorithms that rely on deep neural networks (DNN) are being increasingly employed in workflow automation and processing of large datasets to improve quality assurance and cost-effectiveness. A recent study by Kustner et al. [20] presented a DNN for automatic, reference-free quality assessment in MRI studies of the head, thorax, abdomen, pelvis, as well as whole-body examinations. Used in 2911 datasets obtained from 250 patients, their framework estimated image quality accurately and efficiently. Another study demonstrated good feasibility and accuracy for the automated, reference-free detection of motion artifacts in MRI studies of the head and abdomen via a DNN [21]. These methods are of special interest for retrospective quality control in large cohort studies and, similarly to our study, could provide insight into how the RTs' decision-making correlated with the automated assessments. Their, as well as our, methodologies could also be implemented as prospective quality assurance tools, either by providing in-line quantitative feedback to RTs or by automatically "flagging" relevant acquisitions for repetition, therewith bypassing the need for manual quality assessment altogether, making it reader-independent and consequently experience-independent, and thus speeding up the imaging process. Even automatic adaption of acquisition settings is conceivable. In its current implementation for our study, the computation of quantitative image quality parameters takes too long for this type of application. Deep neural networks likely hold the potential to overcome this hurdle and may consequently pave the way for clinical utilization, which is our long-term goal for this concept of automated image quality assessment.

Our study has certain limitations. We did not examine whether or not professional experience relates to the individual RTs' decision for protocol repetition, due to these data not being available to researchers. Based on the available quantitative data alone, we were also unable to predict which types of artifacts were associated with specific protocol repetitions—further research into their relationship will be necessary, including a visual artifact rating for the initial acquisitions that were subsequently repeated. Although we did examine the effect of image quality on protocol repetition, we did not investigate the

reciprocal effect of protocol repetition on image quality—that is, if repetitions were indeed beneficial to the resulting image data, and in which instances image quality issues may have prevailed (and whether these were of clinical relevance). However, this will be the subject of a follow-up study. Lastly, the assessment of image quality via quantitative parameters was performed as exploratory research and has yet to be validated in a clinical setting with less standardization.

In conclusion, protocol repetitions in MRI are remarkably frequent even in the highly controlled setting of a large cohort study. Automated image quality assessment shows predictive value for the RTs' decision whether or not to perform protocol repetitions, and can objectivize this multifaceted process. Especially when used in conjunction with guided or automated planning tools, it has strong potential to ensure consistent examination quality in prospective MRI studies as well as clinical practice, while simultaneously improving time-efficiency and cost-efficiency.

Data Availability Statement

The data sets generated during and/or analyzed during the current study are not publicly available. However, data are available upon request from NAKO Transferstelle (https://transfer.nako.de/ or transferstelle@nako.de) by means of a project agreement, which will be subject to approval by the NAKO board.

Acknowledgments

This project was conducted with data from the German National Cohort (NAKO; www.nako.de). The NAKO is funded by the Federal Ministry of Education and Research (BMBF) [project funding reference numbers: 01ER1301A/B/C and 01ER1511D], the German federal states, and the Helmholtz Association with additional financial support by the participating universities and the institutes of the Leibniz Association.

The authors thank all members of the German National Cohort MRI Study Investigators **(Appendix)**. The authors are also grateful to all participants who took part in the NAKO study as well as its staff members.

Conflicts of Interest

None declared.

REFERENCES

1. Bild, D.E., et al., *Multi-Ethnic Study of Atherosclerosis: objectives and design.* Am J Epidemiol, 2002. **156**(9): p. 871-81.
2. Mahmood, S.S., et al., *The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective.* Lancet, 2014. **383**(9921): p. 999-1008.
3. Volzke, H., et al., *Cohort profile: the study of health in Pomerania.* Int J Epidemiol, 2011. **40**(2): p. 294-307.
4. Sudlow, C., et al., *UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age.* PLoS Med, 2015. **12**(3): p. e1001779.
5. Bamberg, F., et al., *Whole-Body MR Imaging in the German National Cohort: Rationale, Design, and Technical Background.* Radiology, 2015. **277**(1): p. 206-20.
6. Schlett, C.L., et al., *Population-Based Imaging and Radiomics: Rationale and Perspective of the German National Cohort MRI Study.* Rofo, 2016. **188**(7): p. 652-61.
7. Schmidt, C.O., et al., *Quality standards for epidemiologic cohort studies : An evaluated catalogue of requirements for the conduct and preparation of cohort studies.* Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz, 2018. **61**(1): p. 65-77.
8. Kart, T., et al., *Deep Learning-Based Automated Abdominal Organ Segmentation in the UK Biobank and German National Cohort Magnetic Resonance Imaging Studies.* Invest Radiol, 2021. **56**(6): p. 401-408.
9. Stocker, D., et al., *Performance of an Automated Versus a Manual Whole-Body Magnetic Resonance Imaging Workflow.* Invest Radiol, 2018. **53**(8): p. 463-471.
10. Esser, M., et al., *Performance of an Automated Workflow for Magnetic Resonance Imaging of the Prostate: Comparison With a Manual Workflow.* Invest Radiol, 2020. **55**(5): p. 277-284.
11. Wichmann, H.E., et al., *The German National Cohort.* Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz, 2012. **55**(6-7): p. 781-7.
12. Wood, M.L. and R.M. Henkelman, *MR image artifacts from periodic motion.* Med Phys, 1985. **12**(2): p. 143-51.
13. Huang, L.-K. and M.-J.J. Wang, *Image thresholding by minimizing the measures of fuzziness.* Pattern Recognition, 1995. **28**(1): p. 41-51.
14. Kittler, J. and J. Illingworth, *Minimum error thresholding.* Pattern Recognition, 1986. **19**(1): p. 41-47.
15. Zack, G.W., W.E. Rogers, and S.A. Latt, *Automatic measurement of sister chromatid exchange frequency.* Journal of Histochemistry & Cytochemistry, 1977. **25**(7): p. 741-753.
16. Firbank, M.J., et al., *A comparison of two methods for measuring the signal to noise ratio on MR images.* Physics in Medicine and Biology, 1999. **44**(12): p. N261-N264.
17. Smith, A.M., et al., *Investigation of Low Frequency Drift in fMRI Signal.* NeuroImage, 1999. **9**(5): p. 526-533.
18. Petersen, S.E., et al., *Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of UK Biobank - rationale, challenges and approaches.* J Cardiovasc Magn Reson, 2013. **15**: p. 46.
19. Alfaro-Almagro, F., et al., *Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank.* Neuroimage, 2018. **166**: p. 400-424.
20. Kustner, T., et al., *A machine-learning framework for automatic reference-free quality assessment in MRI.* Magn Reson Imaging, 2018. **53**: p. 134-147.
21. Kustner, T., et al., *Automated reference-free detection of motion artifacts in magnetic resonance images.* Magma, 2018. **31**(2): p. 243-256.

APPENDIX

The German National Cohort MR Imaging Study Investigators

MR Study Site: Augsburg, Germany

Thomas Kroencke, Armin Seifarth, Fabian Bamberg, Wieland Sommer, Maximilian F. Reiser, Birgit Ertl-Wagner, Günter Lewentat, Alexandra Kragler, Alexandra Montag, Erich Wichmann, Jakob Linseisen

MR Study Site: Mannheim/Heidelberg, Germany

Gabriele Ende, Karin Halina Greiser, Tanja Höpker, Andreas Meyer-Lindenberg, Rudolf Kaaks, Hans-Ulrich Kauczor, Heinz-Peter Schlemmer, Christopher L. Schlett, Christopher Schuppert, Ramona Sowade, Oyunbileg von Stackelberg, Sabine Weckbach

MR Study Site: Berlin, Germany

Thoralf Niendorf, Tobias Pischon, Jeanette Schulz-Menger

MR Study Site: Essen, Germany

Michael Forsting, Armin de Greiff, Karl-Heinz Jöckel, Susanne Ladd, Anton Quinston, Börge Schmidt

MR Study Site: Neubrandenburg, Germany

Norbert Hosten, Stefanie Kau, Robin Bülow, Henry Völzke

Imaging Core Coordination and Training Center

Fabian Bamberg, Christopher L. Schlett, Lisa Kretz, Marco Janoschke, Margit Ecker, Michael Klingenberg, Thomas Hendel, Julia Mischner, Michele Picciolo, Maximilian F. Reiser, Sonja Selder

Imaging Core MR Data Management Center

Matthias Günther, Jochen Hirsch, Alexander Köhn, Andreas Thomsen

Imaging Core Incidental Findings Center

Sabine Weckbach, Hans-Ulrich Kauczor, Ramona Sowade, Oyunbileg von Stackelberg

Imaging Core Quality Assurance Centre

Robin Bülow, Norbert Hosten, Stefanie Kau, Henry Völzke

Scientific Focus Group: Musculoskeletal Imaging

Marc-André Weber (chair), Patrick Asbach, Felix Eckstein, Klaus-Peter Günther, Katrin Hegenscheid, Mike Notohamiprodjo, Frank Pessler, Frank Roemer, Armin Seifarth, Christopher L. Schlett, Gerwin Schmidt, Carsten Oliver Schmidt, Sabine Weckbach

Scientific Focus Group: Body Imaging

Hans-Ulrich Kauczor (chair), Fabian Bamberg, Thomas Kröncke, Jong Hee Hwang, Jens-Peter Kühn, Ralf Puls, Michael Roden, Christopher L. Schlett, Johanna Nattenmüller, Ricarda v. Krüchten, Daniel Theisen, Frank Wacker

Scientific Focus Group: Neuroimaging

Michael Forsting (chair), Svenja Caspers (chair), Alfons Schnitzler, Katrin Amunts, Klaus Berger, Susanne Ladd, Arno Villringer, Matthias Günther

Scientific Focus Group: Cardiovascular Imaging

Christopher L. Schlett (chair), Jeanette Schulz-Menger (chair), Marcus Dörr, Tobias Saam, Fabian Bamberg, Henning Steen, Daniel Theisen, Christopher Schuppert, Jana Taron

Central Data Management of the German National Cohort

Daniel Kraft, Rudolf Kaaks, Stefan Ostrzinski, Wolfgang Hoffmann

MR External Advisory Board

Udo Hoffmann (chair), Jeff Carr, Christopher J. O'Donnell, Arfan Ikram, Christopher Kramer, Gabriel Krestin, Steffen Petersen

External Collaboration Partners

Siemens Healthcare (Erlangen, Germany); Joachim Graessner, Regina Hunger, Bernd Ohnesorge, Helder M. Reboredo

FIGURES

Figure 1. The standard acquisition sequence of MRI protocols in the German National Cohort (NAKO) study covers 4 anatomical domains for epidemiologic research in the neurological, cardiovascular, thoracoabdominal, and musculoskeletal fields. The planned scan time without protocol repetitions is approximately 60 minutes.

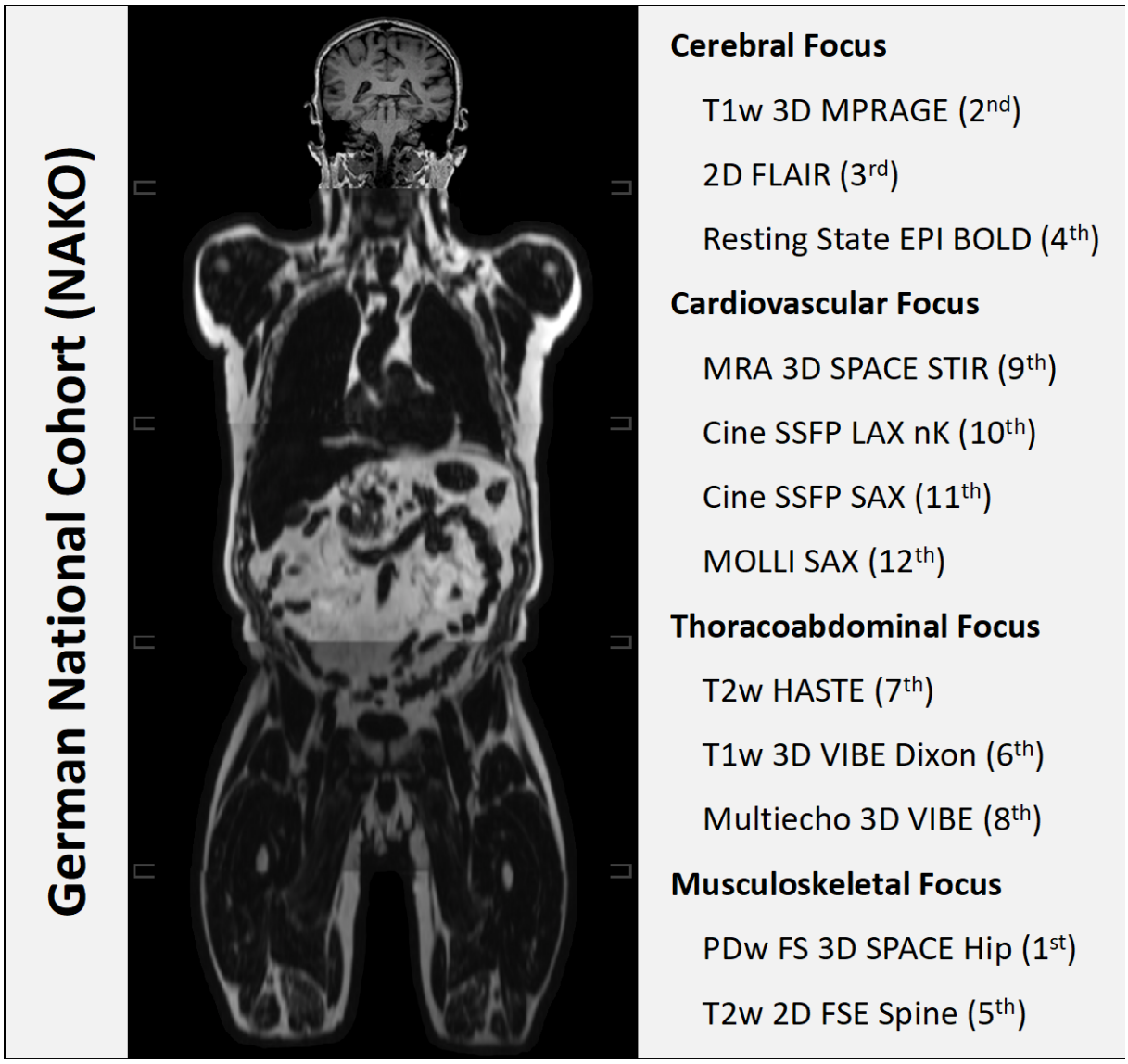


Figure 2. Repetition frequency (here defined as the percentage of participants with at least 1 protocol repetition) varied statistically significantly across sites, as well as over the enrollment period.

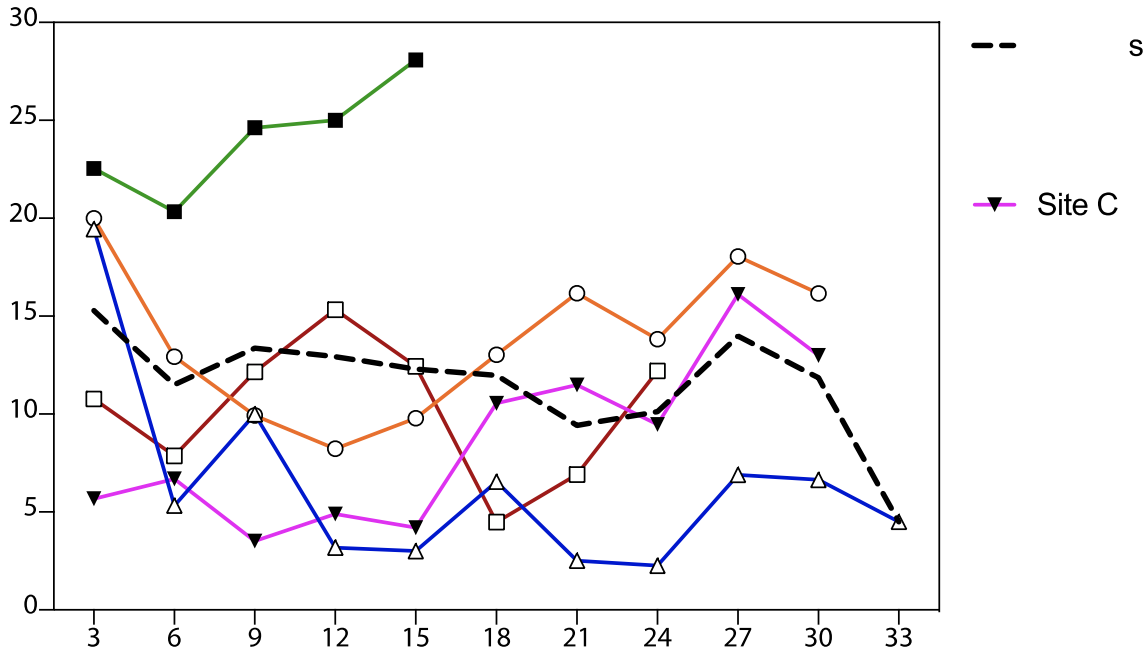


Figure 3. Repetition frequency by protocol across all participants (n = 11,347). As an example, the Neuro 2D FLAIR protocol was repeated at least once in 2.9% of all participants.

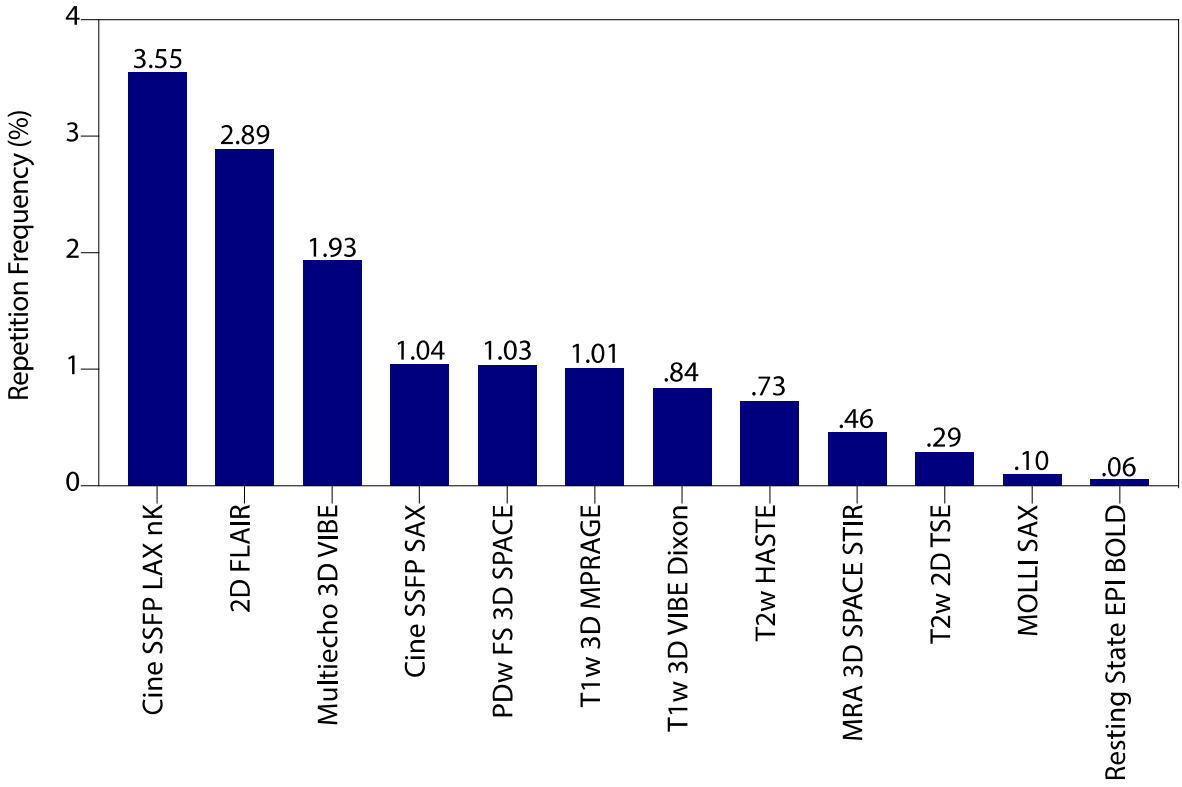


Figure 4. The upper row presents examples of a non-contrast, axial Neuro T1w 3D MPRAGE protocol, acquired according to the German National Cohort (NAKO) study specifications [5]. **A**, Ainitial acquisition of subjectively high quality. **B**, An initial acquisition of subjectively low quality that was subsequently repeated, and a selection of image quality parameters (universal image quality (UQI), signal-to-noise ratio (SNR), and sharpness) confirmed the visual impression. The lower row provides frequency distributions for these parameters in initial acquisitions of the axial Neuro T1w 3D MPRAGE protocol across all participants (further grouped into cases with and without subsequent repetition).

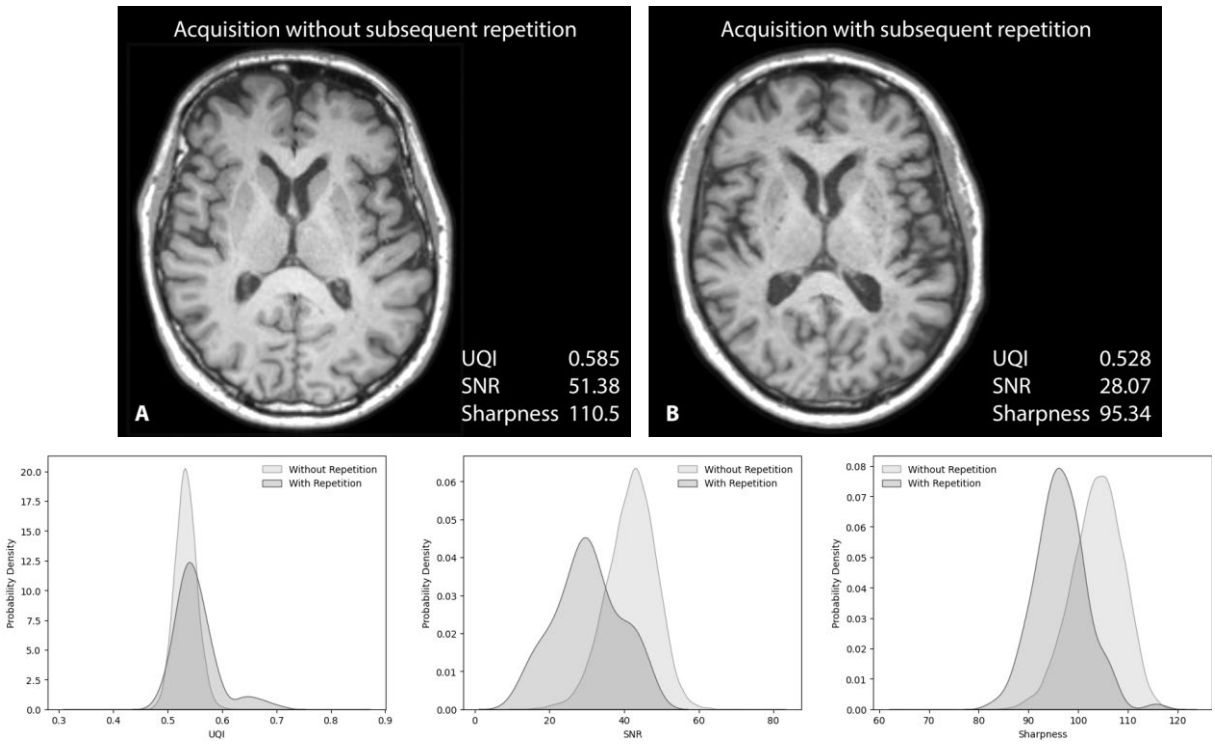


Figure 5. Receiver operating characteristic curves from image quality parameter combinations for each acquired protocol (Cardio Cine SSFP LAX nK represented by 2K). Whichever combination provided the highest area under the curve (AUC) in Supplemental Table 2 was selected for this overview.

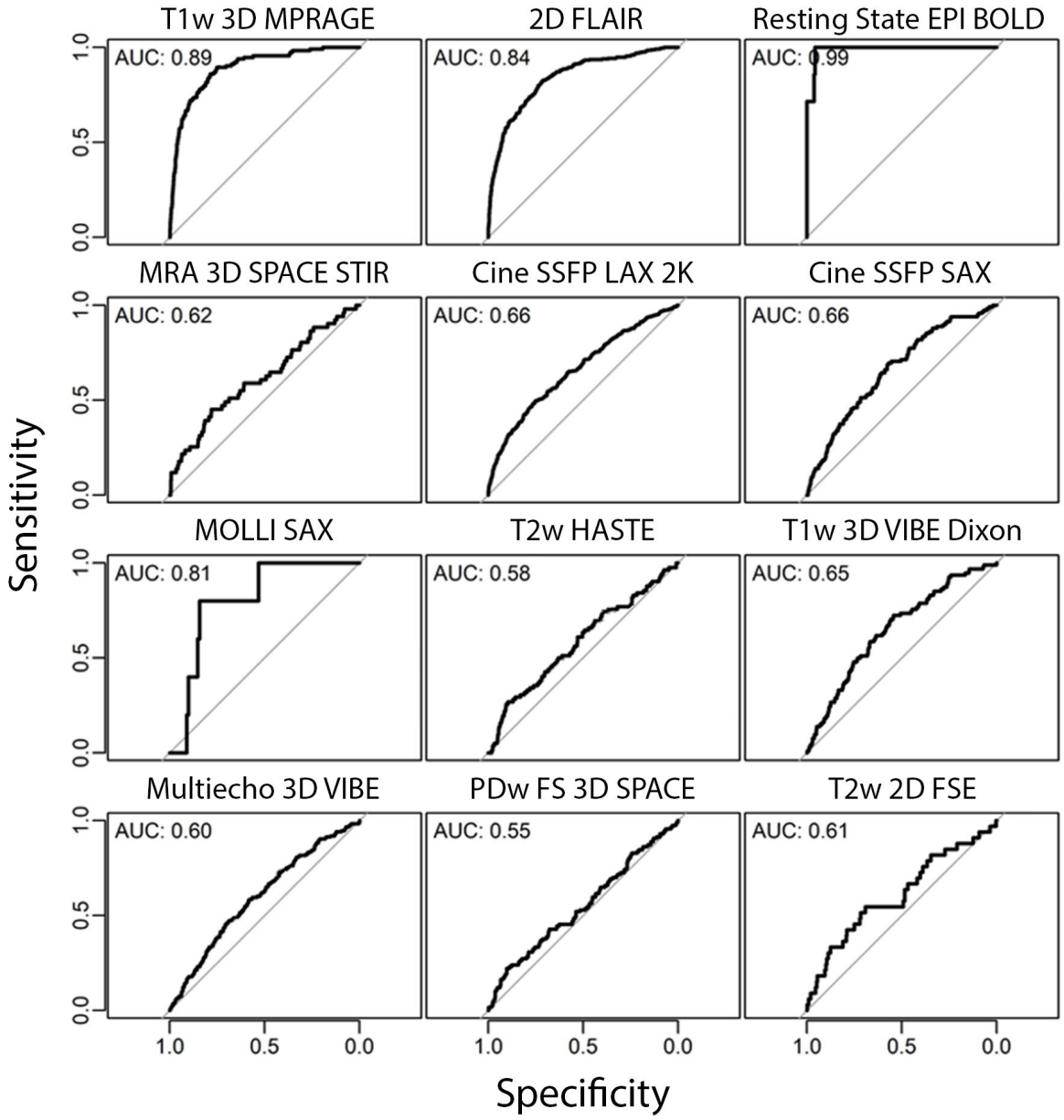
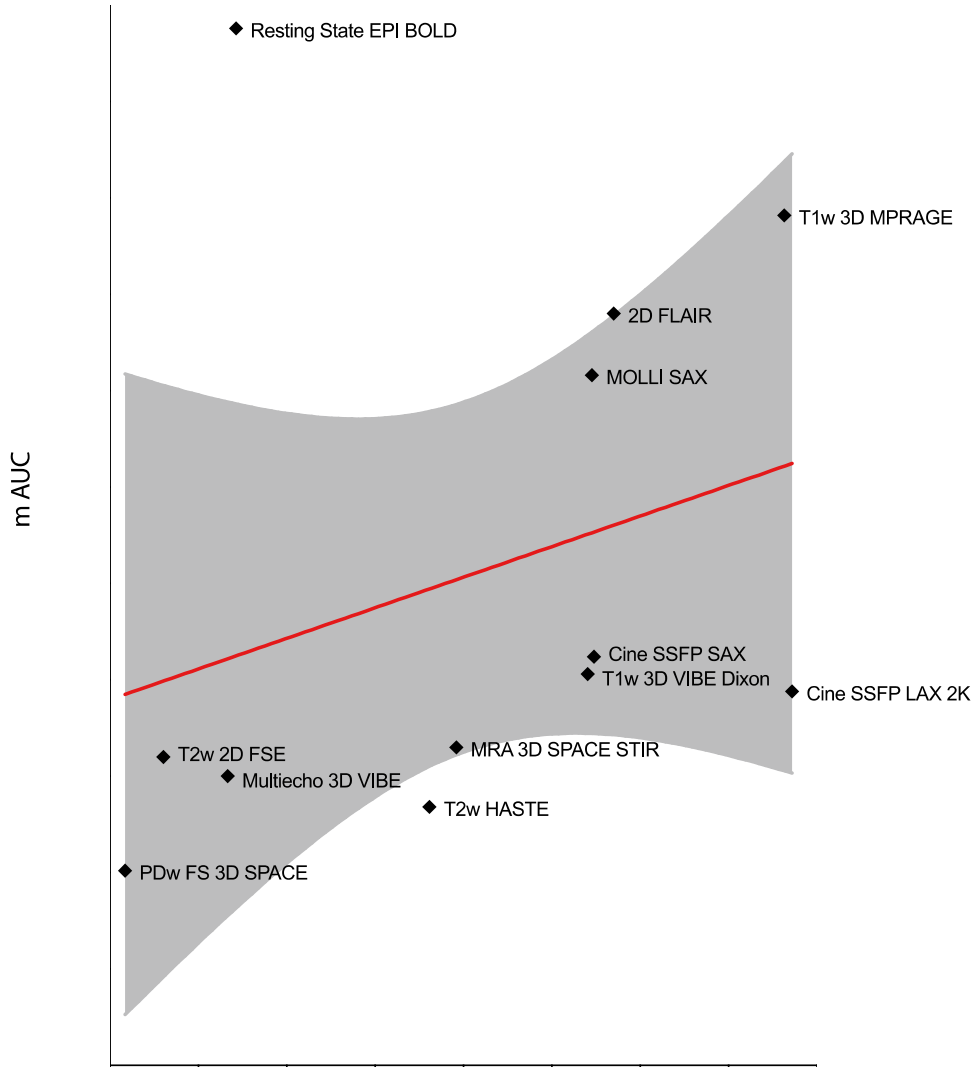


Figure 6. Relationship between manual setup variations and classification ability of image quality parameters. On the *x* axis: proportion of protocol repetitions without manual setup changes of all repetitions (Cardio Cine SSFP LAX-2K representative for nK). On the *y* axis: maximum area under the curve (AUC), given by the best discriminating combination of quality parameters (for initial acquisitions with and without subsequent repetition, cf. Supplemental Tables 2 and 3). Red line: line of linear correlation for all protocols (correlation coefficient: $r = -0.30$). Gray area, 95% confidence interval. In summary, the more often manual setup changes occurred in the context of protocol repetitions, the less were image quality parameters able to discriminate between baseline acquisitions and repetitions.



TABLES

Table 1. Direction of Change of Automatically Derived Image Quality Parameters for Initial Acquisitions With and Without Subsequent Repetition (From “Without Repetition” to “With Repetition”), Along With Statistical Significance and Area Under the Curve.

Protocol	UQI	Sharpness	SNR	Specific SNR	Structured Noise Max	Structured Noise Average	N/2 Ghosting Max	N/2 Ghosting Average	Drift	Variation Over Time	Foreground Ratio
T1w 3D MPRAGE	▲*** 0.65	▼*** 0.85	▼*** 0.85	▼*** 0.68	▲*** 0.73	▲*** 0.76	▲*** 0.66	▲*** 0.64	NA	NA	▲*** 0.68
2D FLAIR	▲*** 0.58	NS	▲*** 0.83	▼*** 0.74	▲*** 0.69	▲*** 0.73	▲*** 0.69	▲*** 0.67	NA	NA	NS
Resting State EPI BOLD	▲*** 0.95	▲*** 0.82	▼*** 0.97	▼*** 0.81	NS	NS	▼* 0.63	▼* 0.69	▼*** 0.72	NS	NS
MRA 3D SPACE STIR	▼** 0.56	▼*** 0.62	NA	NA	NA	NA	NA	NA	NA	NA	NA
Cine SSFP LAX 2K	NS	▼*** 0.57	NS	NA	NS	NS	NS	NS	NA	NA	NS
Cine SSFP LAX 3K	NS	▼*** 0.60	▼*** 0.61	NA	▲*** 0.60	▲*** 0.60	▲* 0.54	▲* 0.54	NA	NA	▲** 0.54
Cine SSFP LAX 4K	NS	▼* 0.53	▼** 0.55	NA	▲*** 0.60	▲*** 0.60	▲* 0.58	▲* 0.58	NA	NA	▼** 0.53
Cine SSFP SAX	NS	NS	▼*** 0.63	NA	▲** 0.59	▲*** 0.62	NS	▲* 0.55	NA	NA	▲** 0.56
MOLLI SAX	NS	NS	NS	NA	NS	NS	NS	NS	NA	NA	NS
T2w HASTE	NS	NS	NS	NA	NS	NS	NS	NS	NA	NA	NS
T1w 3D VIBE Dixon	NS	▼*** 0.64	NS	NA	NS	NS	▼** 0.59	▼** 0.58	NA	NA	▲*** 0.63
Multiecho 3D VIBE	NS	▼*** 0.57	NS	NA	NS	NS	NS	NS	NA	NA	▲*** 0.58
PDw FS 3D SPACE	▲** 0.55	▲* 0.54	NA	NA	NA	NA	NA	NA	NA	NA	NA
T2w 2D FSE	NS	NS	NS	NS	NS	NS	NS	NS	NA	NA	NS

Asterisks denote ranges of *P* values from Student *t* test. Numerical values represent area under the curve.

*$P < 0.05$. **$P < 0.01$. ***$P < 0.001$. NS, not significant; NA, not available; UQI, universal quality index; SNR, signal-to-noise ratio.

532 **Table 2.** Overview of the Technical Adjustments Performed Manually by RTs for Protocol Repetitions (Total N = 1606).

Protocol Repetition	n	%
Without change	**807**	**50.2**
Change in 1 category	**521**	**32.4**
RF coil configuration	8	0.5
Slice position	513	31.9
Slice orientation	0	0
FOV	0	0
Changes in 2 categories	**227**	**14.1**
RF coil configuration + slice position	65	4.0
Slice position + slice orientation	152	9.5
Slice position + FOV	10	0.6
Changes in 3 categories	**51**	**3.2**
Slice Position + slice orientation + RF coil configuration	8	0.5
Slice Position + slice orientation + FOV	43	2.7
Changes in 4 categories	**0**	**0**

533

534 Four categories of change were analyzed: RF coil configuration, slice position, slice orientation, field of view.

535 RF, radiofrequency; FOV, field of view.

536 **SUPPLEMENTAL MATERIAL**

537 **Supplemental Table 1.** Frequency of protocol repetitions and differences in automated

538 image quality parameters.

Protocol Image quality parameter	All initial acquisitions	Initial acquisitions with repetition		Initial acquisitions without repetition		p	AUC
	N	N	Value	N	Value		
T1w 3D MPRAGE (Neuro)							
UQI	11,310	114	0.55	11,196	0.53	**<.001**	0.65
Sharpness	·	·	96.5	·	103.7	**<.001**	0.85
SNR	·	·	30.2	·	41.9	**<.001**	0.85
Specific SNR	·	·	130.7	·	155.4	**<.001**	0.68
Structured Noise Max	·	·	0.69	·	0.63	**<.001**	0.73
Structured Noise Avg	·	·	0.50	·	0.44	**<.001**	0.76
N/2 Ghosting Max	·	·	0.49	·	0.43	**<.001**	0.66
N/2 Ghosting Avg	·	·	0.26	·	0.24	**<.001**	0.64
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,310	114	0.40	11,196	0.38	**<.001**	0.68
2D FLAIR (Neuro)							
UQI	11,294	326	0.42	10,968	0.40	**<.001**	0.58
Sharpness	·	·	71.9	·	71.7	.46	0.59
SNR	·	·	25.2	·	33.7	**<.001**	0.83
Specific SNR	·	·	40.6	·	54.3	**<.001**	0.74
Structured Noise Max	11,291	324	0.74	10,967	0.67	**<.001**	0.69
Structured Noise Avg	·	·	0.55	·	0.49	**<.001**	0.73
N/2 Ghosting Max	·	·	0.55	·	0.46	**<.001**	0.69
N/2 Ghosting Avg	·	·	0.31	·	0.26	**<.001**	0.67
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,294	326	0.62	10,968	0.63	.29	0.52
Resting State EPI BOLD (Neuro)							
UQI	11,266	7	0.57	11,259	0.47	**<.001**	0.95
Sharpness	·	·	106.3	·	94.5	**<.001**	0.82
SNR	·	·	33.9	·	114.9	**<.001**	0.97
Specific SNR	·	·	38.4	·	115.4	**<.001**	0.81
Structured Noise Max	11,251	7	0.581	11,244	0.583	.95	0.49
Structured Noise Avg	·	·	0.50	·	0.54	.33	0.60
N/2 Ghosting Max	11,261	7	0.53	11,254	0.60	**.04**	0.63
N/2 Ghosting Avg	11,262	8	0.42	11,254	0.51	**.04**	0.69
Drift	11,266	7	0.002	11,259	0.006	**<.01**	0.72
Variation Over Time	·	·	0.22	·	0.25	.62	0.62
Foreground Ratio	·	·	0.72	·	0.74	.25	0.45
MRA 3D SPACE STIR (Cardio)							
UQI	11,145	51	0.37	11,094	0.38	**.002**	0.56
Sharpness	·	·	56.7	·	58.9	**<.001**	0.62
SNR	NA	NA	NA	NA	NA	NA	NA

Specific SNR	NA	NA	NA	NA	NA	NA	NA
Structured Noise Max	NA	NA	NA	NA	NA	NA	NA
Structured Noise Avg	NA	NA	NA	NA	NA	NA	NA
N/2 Ghosting Max	NA	NA	NA	NA	NA	NA	NA
N/2 Ghosting Avg	NA	NA	NA	NA	NA	NA	NA
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	NA	NA	NA	NA	NA	NA	NA
Cine SSFP LAX 2K (Cardio)							
UQI	11,054	392	0.294	10,662	0.295	.47	0.51
Sharpness	·	·	43.2	·	43.9	**<.001**	0.57
SNR	·	·	122.6	·	113.7	.17	0.53
Specific SNR	NA	NA	NA	NA	NA	NA	NA
Structured Noise Max	1,758	39	0.79	1,719	0.74	.05	0.64
Structured Noise Avg	·	·	0.79	·	0.74	.05	0.64
N/2 Ghosting Max	1,793	43	0.19	1,750	0.18	.32	0.53
N/2 Ghosting Avg	·	·	0.19	·	0.18	.32	0.53
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,054	392	0.87	10,662	0.86	.35	0.56
Cine SSFP LAX 3K (Cardio)							
UQI	11,053	392	0.282	10,661	0.284	.46	0.52
Sharpness	·	·	49.0	·	50.6	**<.001**	0.60
SNR	·	·	160.0	·	198.9	**<.001**	0.61
Specific SNR	NA	NA	NA	NA	NA	NA	NA
Structured Noise Max	10,407	350	0.69	10,057	0.65	**<.001**	0.60
Structured Noise Avg	·	·	0.69	·	0.65	**<.001**	0.60
N/2 Ghosting Max	10,503	352	0.22	10,151	0.21	**.04**	0.54
N/2 Ghosting Avg	·	·	0.22	·	0.21	**.04**	0.54
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,053	392	0.75	10,661	0.73	**.004**	0.54
Cine SSFP LAX 4K (Cardio)							
UQI	11,053	392	0.2787	10,661	0.2786	.95	0.50
Sharpness	·	·	46.4	·	46.8	**.049**	0.53
SNR	·	·	264.6	·	296.7	**.003**	0.55
Specific SNR	NA	NA	NA	NA	NA	NA	NA
Structured Noise Max	2,623	110	0.64	2,513	0.59	**<.001**	0.60
Structured Noise Avg	·	·	0.64	·	0.59	**<.001**	0.60
N/2 Ghosting Max	2,757	119	0.16	2,638	0.14	**.010**	0.58
N/2 Ghosting Avg	·	·	0.16	·	0.14	**.010**	0.58
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,053	392	0.87	10,661	0.88	**.001**	0.53
Cine SSFP SAX (Cardio)							
UQI	11,020	115	0.312	10,905	0.306	.06	0.52
Sharpness	·	·	49.0	·	49.6	.08	0.57
SNR	·	·	116.3	·	152.0	**<.001**	0.63
Specific SNR	NA	NA	NA	NA	NA	NA	NA
Structured Noise Max	10,686	107	0.82	10,579	0.79	**.003**	0.59
Structured Noise Avg	·	·	0.72	·	0.68	**<.001**	0.62

N/2 Ghosting Max	10,715	107	0.275	10,608	0.266	.31	0.52
N/2 Ghosting Avg	·	·	0.17	·	0.16	**.026**	0.55
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,020	115	0.80	10,905	0.78	**.009**	0.56
MOLLI SAX (Cardio)							
UQI	11,022	11	0.269	11,011	0.267	.90	0.54
Sharpness	·	·	42.6	·	43.5	.55	0.57
SNR	·	·	21.1	·	24.5	.66	0.56
Specific SNR	NA	NA	NA	NA	NA	NA	NA
Structured Noise Max	5,137	7	0.91	5,130	0.85	.20	0.67
Structured Noise Avg	·	·	0.91	·	0.85	.20	0.67
N/2 Ghosting Max	4,145	5	0.14	4,140	0.18	.31	0.63
N/2 Ghosting Avg	·	·	0.14	·	0.18	.31	0.63
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,022	11	0.51	11,011	0.57	.49	0.59
T2w HASTE (Body)							
UQI	11,182	82	0.20	11,100	0.21	.47	0.53
Sharpness	·	·	61.0	·	62.4	.23	0.55
SNR	·	·	129.1	·	131.4	.57	0.52
Specific SNR	NA	NA	NA	NA	NA	NA	NA
Structured Noise Max	11,182	82	0.6828	11,100	0.6832	.94	0.50
Structured Noise Avg	·	·	0.465	·	0.470	.27	0.54
N/2 Ghosting Max	·	·	0.35	·	0.36	.25	0.55
N/2 Ghosting Avg	·	·	0.22	·	0.23	.76	0.50
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,182	82	0.56	11,100	0.54	.06	0.56
T1w 3D VIBE Dixon (Body)							
UQI	11,194	94	0.35	11,000	0.36	.52	0.53
Sharpness	·	·	71.6	·	74.1	**<.001**	0.64
SNR	·	·	74.9	·	78.7	.084	0.55
Specific SNR	·	·	NA	·	NA	NA	NA
Structured Noise Max	·	·	0.66	·	0.67	.20	0.55
Structured Noise Avg	·	·	0.43	·	0.42	.25	0.53
N/2 Ghosting Max	·	·	0.37	·	0.41	**.001**	0.59
N/2 Ghosting Avg	·	·	0.20	·	0.22	**.005**	0.58
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,194	94	0.56	11,000	0.52	**<.001**	0.63
Multiecho 3D VIBE (Body)							
UQI	11,174	216	0.237	10,958	0.236	.73	0.53
Sharpness	·	·	57.3	·	59.1	**<.001**	0.57
SNR	·	·	115.8	·	113.6	.30	0.52
Specific SNR	NA	NA	NA	NA	NA	NA	NA
Structured Noise Max	9,836	186	0.577	9,650	0.577	.97	0.50
Structured Noise Avg	·	·	0.466	·	0.462	.50	0.52
N/2 Ghosting Max	9,869	186	0.296	9,683	0.290	.41	0.53
N/2 Ghosting Avg	·	·	0.191	·	0.187	.49	0.53
Drift	NA	NA	NA	NA	NA	NA	NA

Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,174	216	0.77	10,958	0.74	**<.001**	0.58
PDw FS 3D SPACE (MSK)							
UQI	11,346	117	0.314	11,229	0.306	**.008**	0.55
Sharpness	·	·	52.3	·	51.6	**.022**	0.54
SNR	NA	NA	NA	NA	NA	NA	NA
Specific SNR	NA	NA	NA	NA	NA	NA	NA
Structured Noise Max	NA	NA	NA	NA	NA	NA	NA
Structured Noise Avg	NA	NA	NA	NA	NA	NA	NA
N/2 Ghosting Max	NA	NA	NA	NA	NA	NA	NA
N/2 Ghosting Avg	NA	NA	NA	NA	NA	NA	NA
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	NA	NA	NA	NA	NA	NA	NA
T2w 2D FSE (MSK)							
UQI	11,232	33	0.354	11,199	0.351	.71	0.53
Sharpness	·	·	68.5	·	67.8	.37	0.55
SNR	·	·	28.8	·	29.1	.88	0.52
Specific SNR	·	·	335.5	·	353.1	.40	0.54
Structured Noise Max	·	·	0.47	·	0.46	.50	0.51
Structured Noise Avg	·	·	0.451	·	0.449	.78	0.49
N/2 Ghosting Max	·	·	0.18	·	0.19	.75	0.48
N/2 Ghosting Avg	·	·	0.162	·	0.164	.68	0.48
Drift	NA	NA	NA	NA	NA	NA	NA
Variation Over Time	NA	NA	NA	NA	NA	NA	NA
Foreground Ratio	11,232	33	0.406	11,199	0.415	.36	0.55

Note: Centered dot denotes the same value as above. Probability values from Student's t-test. Bold denotes statistical significance at level p<.05. AUC: area under the curve.

Supplemental Table 2. Discrimination ability of combined image quality parameters.

Protocol	Model Number										
	1	2	3	4	5	6	7	8	9	10	11
T1w 3D MPRAGE	0.88	0.88	0.88	0.87	0.87	0.89	**0.89**	0.85	0.73	0.84	0.85
2D FLAIR	0.66	0.84	**0.84**	0.83	0.83	0.84	0.51	0.84	0.79	0.73	0.83
Resting State EPI BOLD	0.82	0.97	0.97	0.97	0.96	0.95	0.77	0.97	0.81	0.88	**0.99**
MRA 3D SPACE STIR	**0.62**	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Cine SSFP LAX 2K	0.61	0.61	NA	NA	0.57	NA	0.57	0.54	NA	0.63	0.63
Cine SSFP LAX 3K	0.64	0.66	NA	NA	0.65	NA	0.60	0.64	NA	0.64	0.62
Cine SSFP LAX 4K	0.55	0.56	NA	NA	0.56	NA	0.56	0.55	NA	0.60	0.60
Cine SSFP SAX	0.65	**0.66**	NA	NA	0.63	NA	0.56	0.66	NA	0.62	0.65
MOLLI SAX	0.59	0.58	NA	NA	0.58	NA	0.57	0.59	NA	0.69	0.62
T2w HASTE	0.55	0.54	NA	NA	0.54	NA	0.57	0.56	NA	0.57	0.55
T1w 3D VIBE Dixon	0.64	0.65	NA	NA	0.64	NA	0.64	0.64	NA	0.64	0.60
Multiecho 3D VIBE	0.58	0.59	NA	NA	0.59	NA	0.58	0.58	NA	0.59	0.54
PDw FS 3D SPACE	**0.55**	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T2w 2D FSE	0.54	0.54	0.59	0.58	0.55	0.57	0.55	0.55	0.55	0.59	0.58

Protocol	Model Number (continued)											
	12	13	14	15	16	17	18	19	20	21	22	23
T1w 3D MPRAGE	0.87	0.85	0.85	0.88	0.76	0.66	0.73	0.79	0.79	0.88	0.81	NA
2D FLAIR	0.84	0.70	0.83	0.84	0.72	0.69	0.69	0.73	0.72	0.84	0.76	NA
Resting State EPI BOLD	0.99	0.85	0.98	0.97	0.80	0.65	0.72	0.69	0.82	0.99	0.92	0.98
MRA 3D SPACE STIR	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Cine SSFP LAX 2K	0.62	0.57	0.58	0.59	0.64	0.53	**0.64**	**0.64**	**0.64**	0.63	NA	NA
Cine SSFP LAX 3K	0.66	0.61	0.61	0.65	0.60	0.54	0.61	0.61	0.61	**0.66**	NA	NA
Cine SSFP LAX 4K	0.61	0.58	0.60	0.61	0.60	0.58	0.60	0.60	0.60	**0.61**	NA	NA
Cine SSFP SAX	0.65	0.60	0.64	0.65	0.62	0.56	0.59	0.62	0.62	0.65	NA	NA
MOLLI SAX	0.71	0.69	**0.81**	0.81	0.67	0.63	0.77	0.77	0.77	0.81	NA	NA
T2w HASTE	0.57	0.56	0.56	0.56	0.54	0.56	0.54	0.54	0.58	**0.58**	NA	NA
T1w 3D VIBE Dixon	0.65	0.64	0.62	0.65	0.58	0.60	0.59	0.60	0.61	**0.65**	NA	NA
Multiecho 3D VIBE	0.59	0.58	0.53	0.60	0.53	0.53	0.53	0.52	0.53	**0.60**	NA	NA
PDw FS 3D SPACE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
T2w 2D FSE	0.59	0.56	0.51	0.56	**0.61**	0.52	0.48	0.45	0.56	0.55	0.56	NA

Note: Predictors are composed of parameter combinations as specified below. AUC values are shown for "initial acquisition with subsequent repetition" versus "initial acquisition without subsequent repetition". Bold marks the maximum area under the curve (AUC) value for each protocol.

Legend for model numbers and predictors:

1	UQI + Sharpness
2	UQI + Sharpness + SNR
3	UQI + Sharpness + SNR + Specific SNR
4	Sharpness + SNR + Specific SNR
5	Sharpness + SNR
6	Sharpness + SNR + Specific SNR + Foreground Ratio
7	Sharpness + Foreground Ratio
8	SNR + Foreground Ratio
9	Specific SNR + Foreground Ratio
10	Sharpness + Structured Noise Max + Structured Noise Avg
11	SNR + Structured Noise Max + Structured Noise Avg
12	Sharpness + SNR + Structured Noise Max + Structured Noise Avg
13	Sharpness + N/2 Ghosting Max + N/2 Ghosting Avg
14	SNR + N/2 Ghosting Max + N/2 Ghosting Avg
15	Sharpness + SNR + N/2 Ghosting Max + N/2 Ghosting Avg
16	Structured Noise Max + Structured Noise Avg
17	N/2 Ghosting Max + N/2 Ghosting Avg
18	Structured Noise Max + N/2 Ghosting Max
19	Structured Noise Avg + N/2 Ghosting Avg
20	Structured Noise Max + Structured Noise Avg + N/2 Ghosting Max + N/2 Ghosting Avg
21	Sharpness + SNR + Structured Noise Max + Structured Noise Avg + N/2 Ghosting Max + N/2 Ghosting Avg
22	Specific SNR + Structured Noise Max + Structured Noise Avg + N/2 Ghosting Max + N/2 Ghosting Avg
23	UQI + Sharpness + SNR + Specific SNR + Structured Noise Max + Structured Noise Avg + N/2 Ghosting Max + N/2 Ghosting Avg + Drift + Variation Over Time + Foreground Ratio

Supplemental Table 3. Detailed analysis of the technical adjustments performed manually by the radiologic technologist for protocol repetitions (including any first, and, if acquired, second and third repetition).

Protocol	Repetitions	No Change		1 Change				2 Changes						3 Changes			
				Coil Conf.		Slice Pos.		Coil Conf. Slice Pos.		FOV Slice Pos.		Slice Pos. Slice Orient.		Coil Conf. Slice Pos. Slice Orient.		FOV Slice Pos. Slice Orient.	
	N	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)
All	1,606	807	(50.2)	8	(0.5)	513	(31.9)	65	(4.0)	10	(0.6)	152	(9.5)	8	(0.5)	43	(2.7)
T1w 3D MPRAGE	114	87	(76.3)	4	(3.5)	16	(14)	2	(1.8)	0		3	(2.6)	2	(1.8)	0	
2D FLAIR	335	191	(57.0)	0		45	(13.4)	1	(0.3)	0		96	(28.7)	2	(0.6)	0	
Resting State EPI BOLD	7	1	(14.3)	0		1	(14.3)	0		0		4	(57.1)	1	(14.3)	0	
MRA 3D SPACE STIR	51	20	(39.2)	0		31	(60.8)	0		0		0		0		0	
Cine SSFP LAX nK	417	322	(77.2)	0		41	(9.8)	0		2	(0.5)	43	(10.3)	0		9	(2.2)
Cine SSFP SAX	115	63	(54.8)	0		44	(38.3)	0		8	(7.0)	0		0		0	
MOLLI SAX	11	6	(54.5)	0		1	(9.1)	0		0		0		0		4	(36.4)
T2w HASTE	83	30	(36.1)	2	(2.4)	35	(42.2)	15	(18.1)	0		0		1	(1.2)	0	
T1w 3D VIBE Dixon	98	53	(54.1)	1	(1.0)	43	(43.9)	1	(1.0)	0		0		0		0	
Multiecho 3D VIBE	225	30	(13.3)	0		189	(84.0)	6	(2.7)	0		0		0		0	
PDw FS 3D SPACE	117	2	(1.7)	1	(0.9)	66	(56.4)	40	(34.2)	0		6	(5.1)	2	(1.7)	0	
T2w 2D FSE	33	2	(6.1)	0		1	(3.0)	0		0		0		0		30	(90.9)

Note: Four categories of change were analyzed: RF Coil Configuration, Slice Position, Slice Orientation, field of view (FOV). If more than one repetition was performed, changes were evaluated against the initial acquisition.