# BALANCED STEADY-STATE FREE PRECESSION IMAGING OF THE TEMPORAL BONE AND PARANASAL SINUSES AT 0.5T

by

Sarah Reeve

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
January 2023

# Table of Contents

# List of Tables

## List of Figures

## Abstract

Magnetic resonance imaging (MRI) performed with main magnetic field strengths below the conventional 1.5T or 3T has typically been considered inferior, due, primarily, to the proportional decrease in raw signal available. However, recent advancements in gradient systems and radiofrequency receive chains have opened the door to new image acquisition strategies that benefit from use at low field. Hence, this thesis investigates balanced steady-state free precession (bSSFP) protocols designed for a 0.5T system with fast, strong gradients. The protocols are used to image regions that demonstrate susceptibility-induced off-resonance effects, which require phase-cycling techniques to mitigate banding artifacts at conventional field strengths.

The work presented herein consists of three studies. The first, titled "*Artifact-resistant balanced steady-state free precession imaging of the temporal bone and paranasal sinuses without phase-cycling at 0.5T*" illustrates the high artifact tolerance and signal-to-noise ratio attained, at clinical resolutions, by the 0.5T bSSFP protocols. While these metrics are informative, the true value of an image is based on a radiologist's ability to use it to answer a clinical question. To that end, the second study, titled "*Low-field vs. conventional field balanced steady-state free precession imaging of the temporal bone: radiologist rating of anatomical visualization*" examines radiologists' ability to visualize structures of the temporal bone when viewing images acquired with bSSFP at 0.5T or phase-cycled bSSFP at 3T. Analyses revealed no significant difference, overall, in radiologists' ratings, indicating the images were of similar quality. Finally, the third study, titled "*Super resolution allows for the benefits of low resolution balanced steady-state free precession imaging without degrading image quality*", applies a machine learning pipeline to low resolution images resulting from a faster 0.5T bSSFP temporal bone acquisition, bringing them back to high resolution. Evaluation of image ratings shows that radiologists' ability to visualize temporal bone structures is not significantly reduced, thereby permitting a reduction in scan time.

Together, these works illustrate the new opportunities afforded by low-field bSSFP imaging with high-performance gradient systems. These advantages, combined with super resolution techniques, have the potential to make MRI of challenging regions fast and artifact-free, thereby improving patient experience without sacrificing image utility.

## List of Symbols and Abbreviations Used

| | |
|---|---|
| $B_0$ | Main magnetic field strength |
| $\vec{B}$ | Magnetic field vector |
| $T_1$ | Longitudinal relaxation constant |
| $T_2$ | Transverse relaxation constant for relaxation due to atomic and molecular disturbances) |
| $T_2'$ | Transverse relaxation constant for relaxation due to field inhomogeneities |
| $T_2^*$ | Transverse relaxation constant |
| $\vec{M}$ | Net magnetization vector |
| $\vec{B}$ | Magnetic field vector |
| $M_0$ | Equilibrium magnetization magnitude |
| $\gamma$ | Gyromagnetic ratio |
| $\vec{M}_{xy}$ | Transverse magnetization vector |
| $\alpha$ | Flip angle |
| $\omega$ | Larmour frequency |
| $\vec{G}$ | Gradient field vector |
| $\Phi$ | Phase |
| $\rho$ | Effective spin density |
| $\vec{k}$ | K-space vector |
| t | Time |
| s | MRI signal |
| $\beta$ | Effective flip angle |
| $\nu$ | Angular frequency |
| $HR_{Int.}$ | High resolution acquisition interpolated to clinical resolution |
| $HR_{Sr}$ | High resolution acquisition passed through a super resolution pipeline |
| $LR_{Int.}$ | Low resolution acquisition interpolated to clinical resolution |
| $LR_{SrInt.}$ | Low resolution acquisition passed through a super resolution pipeline and subsequent interpolation |
| $W_{PSNR}$ | Set of "PSNR-driven" model weights |
| $W_{NC}$ | Set of "Noise-cancelling" model weights |
| $W_{Interp.}$ | Set of model weights interpolated between $W_{PSNR}$ and $W_{NC}$ |
| $\lambda$ | Weighting factor for $W_{Interp.}$ |

| | |
|---|---|
| bSSFP | Balanced Steady-State Free Precession |
| CNN | Convolutional Neural Network |
| CSF | Cerebral Spinal Fluid |
| CT | Computed Tomography |
| EPI | Echo Planar Imaging |
| FID | Free Induction Decay |
| FIESTA-C | Fast Imaging Employing Steady-State Acquisition - Constructive |
| FOV | Field of View |
| FSE | Fast Spin Echo |
| GAN | Generative Adversarial Network |
| GRAPPA | Generalized Autocalibrating Partial Parallel Acquisition |
| GRE | Gradient Echo |
| HR | High Resolution |
| IAC | Internal Auditory Canal |
| LR | Low Resolution |
| ML | Machine Learning |
| NEX | Number of Excitations |
| MIP | Maximum Intensity Projection |
| MRI | Magnetic Resonance Imaging |
| MSE | Mean Squared Error |
| PDF | Probability Density Function |
| PNS | Peripheral Nerve Stimulation |
| PSNR | Peak Signal-to-Noise Ratio |
| REB | Research Ethics Board |
| ResNet | Residual Network |
| ReLU | Rectified Linear Unit |
| RF | Radio Frequency |
| ROI | Region Of Interest |
| SAR | Specific Absorption Rate |
| SENSE | Sensitivity Encoding |
| SNR | Signal-to-Noise Ratio |
| SOC | Standard Of Care |
| SR | Super Resolution |
| SRCNN | Super Resolution Convolutional Neural Network |
| SRGAN | Super Resolution Generative Adversarial Network |
| SSC | Superior Semicircular Canal |
| SSFP | Steady-State Free Precession |
| TE | Echo Time |
| TR | Repetition Time |

## Acknowledgements

The past four years have been challenging for many reasons, and I must acknowledge a number of people that have made this work possible.

First, my supervisor, Dr. Steven Beyea, and committee members, Dr. Chris Bowen, Dr. James Rioux, and Dr. Alessandro Guida, have each provided invaluable guidance, support, and incredible amounts of knowledge. I cannot thank them enough for creating a welcoming space, literally and metaphorically, that has helped me grow as a scientist and a person. My lab mates, Michelle Pryde and Ally Klassen, were approachable sounding boards and kind friends, without whom my years of research would not have been as productive nor as pleasant. The rest of the BIOTIC team, including Beverly Lieuwen, helped things run smoothly, and provided meaningful input that cannot go unrecognized. There are many others, including Mark Parker, multiple radiologists (Dr. David Volders, Dr. Elena Adela Cora, Dr. Matthias Schmidt, and Dr. David Morris), radiology fellows (Dr. Undrakh-Erdene Erdenebold and Dr. Brian King), MR technologists, and those at the Dalhousie University Medical Physics and Physics departments that provided essential contributions to this work. It has been a great pleasure to be a part of this research community.

My classmates in Medical Physics have been both colleagues and friends, and I am truly grateful for their constant understanding and insight. I wouldn't be here without them.

Finally, I must thank those closest to me. My brilliant and kind sister, Catherine, provided love, perspective, and wisdom that made the last few years possible, and my partner, Jon, made me laugh even on the toughest days. Thank you for being by my side.

# Chapter 1

# Introduction

## 1.1 Preamble

Magnetic resonance imaging (MRI) is a non-invasive, versatile imaging modality with unsurpassed soft tissue contrast, and it is an integral tool in contemporary clinical medicine.[1]

Advancements in MR system components including magnets, gradient systems, and radiofrequency (RF) receive chain technology have contributed to renewed interest in MR imaging at field strengths below the conventional 1.5T, opening the door to new acquisition strategies. Balanced steady-state free precession (bSSFP) is one such acquisition strategy that is now seeing increased application on low-field systems, as it demonstrates high signal-to-noise ratio (SNR) efficiency - important when dealing with decreased polarization of spins at low fields - and enables high resolution imaging in clinically acceptable scan times. The increase in field homogeneity, and decrease in field distortions caused by magnetic susceptibility differences at low field, can minimize artifacts due to off-resonance effects, common to bSSFP acquisitions. These advantages suggest that effective low-field bSSFP imaging could be performed for challenging regions with multiple air/tissue and air/bone interfaces such as the temporal bone and paranasal sinuses.

In addition to hardware developments, the application of machine learning techniques to MR image reconstruction has further modified the landscape of what was once possible with low-field MR acquisitions. Super resolution algorithms, for example, are capable of generating high resolution images from faster, low resolution scans, permitting a wider range of acquisition parameters and lessening the gaps in the well-known trade-off between SNR, resolution, and acquisition time in MRI.

The overall goal of the research presented in this thesis is to investigate the potential afforded by low-field bSSFP imaging and super resolution techniques in the context of temporal bone and paranasal sinus acquisitions. To that end, this work

begins by examining the artifact tolerance of bSSFP protocols designed to image the temporal bone and paranasal sinuses at 0.5T. In the second study, the results of the temporal bone image evaluation are put into a clinical context by employing board-certified neuroradiologists to rate their ability to visualize clinically relevant structures. Finally, the ability to reap the benefits of low resolution bSSFP acquisitions is explored by application of a super resolution pipeline.

## 1.2 History

### 1.2.1 Field Strength

MRI relies on the polarization of hydrogen nuclei (or "spins") in the body, accomplished by placing the subject in a strong, static magnetic field of strength $B_0$ (see Chapter 2 for a comprehensive description of signal generation and image reconstruction). The earliest clinically useful MR images of humans were acquired in 1980 using a resistive magnet that created a static magnetic field with $B_0 = 0.014$ T.[2] The introduction of superconducting magnets to MR imaging in the early 1980s made it possible to generate much stronger fields, and the first commercial units were produced with a field strength of 0.35 T.[2,3] The field strength of choice rapidly increased to 1 T in the mid 1980s and by the 1990s, superconducting systems with $B_0 = 1.5$T were the clinical standard. 3T systems were available by the early 2000s, but were slow to be used clinically, with motion artifacts being difficult to remedy at the high resolutions achieved. Over time, however, 3T became seen as superior to 1.5T for particular applications. By 2020, systems with field strengths equal or greater than 1.5 T represented approximately 85% of market size in North America and Europe.[4] Today, systems operating at 7 T have also seen clinical use for specific applications.[5]

This march towards ever higher field strengths has been motivated, primarily, by the proportional increase in the number of spins that become polarized, which are required to generate an MR signal. The raw signal obtained can then be used to generate images with higher and higher SNR values. Historically, attaining the highest number of polarized spins possible was considered crucial, as conventional MRI required a trade-off between SNR, spatial resolution, and scan time, as demonstrated in Figure 1.1. For example, a common approach to acquiring an image with sufficient

resolution and SNR is to average the data acquired over multiple measurements (i.e. number of excitations, NEX), increasing imaging time by an equivalent factor.

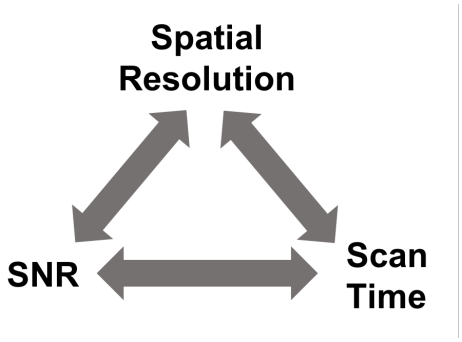


Figure 1.1: The trade-off between SNR, spatial resolution, and scan time in MRI.

However, a comparison of images acquired on early and modern 0.5 T MR systems (Figure 1.2) clearly demonstrates that field strength - and therefore the number of polarized spins available for imaging - is not the limiting factor in achieving diagnostic quality images.[2]

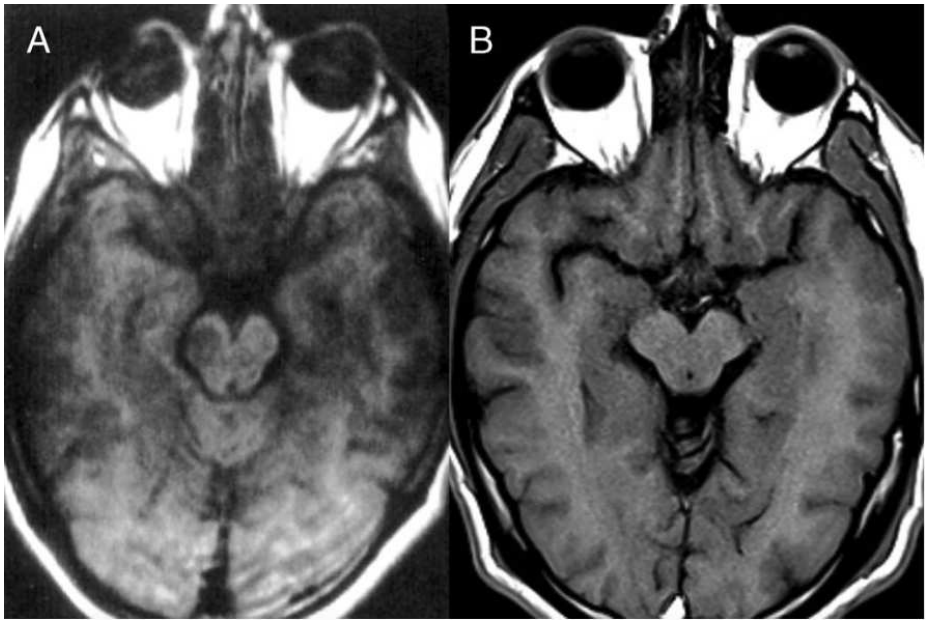


Figure 1.2: MRI of the brain acquired at $B_0 = 0.5$ T in (a) 1984 and (b) 2020. Despite identical field strengths, the image in (B) was acquired in under half the time, and demonstrates improved SNR, contrast, and resolution. Figure is reproduced from Runge et al. 2020.[2]

### 1.2.2 Gradients and RF Receive Coils

Ancillary system components, such as gradient systems, and RF receive chain technology have also seen significant improvements since the first clinical MR systems.

Importantly, data acquisition in MRI does not take place in image space. Instead, information is collected in the spatial frequency domain and placed into a data array known as k-space (see Section 2.1.2), such that an image and its k-space data are related by Fourier transformation. Linearly varying weak magnetic fields, or gradient fields, are utilized in addition to the static field to systematically encode the spatial frequency information required across the volume to be imaged. The maximum gradient strength used, as well as the time required to reach it (i.e. the *rise time*), dictate the maximum attainable resolution for a given scan time, or alternatively, the scan time for a given resolution. Since clinical MRI began in the 1980s, gradient strengths have increased from 1 mT/m to 300 mT/m, with possible *slew rates* (the ratio of maximum gradient strength to rise time) increasing from 3 T/m/s to 400 T/m/s.[6] Importantly, the physiological limit of peripheral nerve stimulation (PNS) prevents the use of high slew rates on whole body systems. However, some modern head-only systems use asymmetric gradient designs,[7,8] which allow the patient's head to be at the isocenter of the gradient field while minimizing the field experienced by the rest of the body. This design reduces PNS concerns significantly and permits the use of high slew rates, though acquisition strategies must still consider patient heating restrictions.

Radio frequency coils are a vital part of the MR system. Transmit coils emit magnetic flux at radio frequencies, causing polarized spins to tip away from the $z$-axis and into the transverse plane, generating a measurable signal. This signal is then captured by the RF receive coils.[9] The receive coil on the first whole-body MR system consisted of wires wound around the bore of the magnet,[10] detecting the signal from the region of interest as well as noise emanating from anywhere within its diameter. Small surface coils, designed to fit closely to the anatomy being examined, were developed in the 1980s and significantly reduced the measurement of unwanted noise from outside the region of interest,[11] improving the SNR of the resulting images. Switchable arrays of coils, first used in the early 1990s, permitted high-sensitivity imaging of large regions (e.g. the spine) without the need to adjust the patient.

However, such arrays still made use of a single receive channel, and "switching" of the coil being used to image was therefore required. Switchable arrays paved the way for the use of multiple coils with independent receive chains, known as phase-array coils, again improving SNR significantly. Today, multi-channel RF receive coils can be made to adjust or conform to the anatomy to be imaged, enabling patient-specific placement and improved measurement efficiency.

### 1.2.3 Data Acquisition and Reconstruction

The hardware developments discussed in the previous section have decreased acquisition times significantly to that considered comfortable for most patients to endure ($\approx$ 5 minutes), while maintaining image quality. However, data acquisition and reconstruction techniques have improved in parallel to hardware advancements and permit accelerated acquisitions that do not necessarily rely on gradient performance.

Partial Fourier techniques, introduced in 1986,[12] require as little as half of k-space to be acquired, and rely on the symmetry of k-space to estimate the remaining portion and produce an image. In practice, field inhomogeneities, susceptibility effects, patient motion, etc. can cause mistakes in spatial encoding, and effort must be taken to correct these errors prior to reconstruction.

Parallel imaging emerged in the 1990s, following the development of phased-array coils. Parallel imaging uses knowledge of the sensitivity profile of each individual coil to assist in spatial localization, allowing for a systematic reduction in the number of k-space points acquired, and a corresponding reduction in scan time. A Fourier transform of each incomplete k-space would result in images that exhibit aliasing artifacts (see Section 2.1.3), however methods were developed to reconstruct a single, un-aliased image, which operate in the spatial frequency domain (e.g. GeneRalized Autocalibrating Partial Parallel Acquisition, GRAPPA[13]) or in the image domain (e.g. SENSitivity Encoding, SENSE[14]). These techniques provided the greatest gain in imaging speed since the conception of MRI,[15] but come at the cost of reduced SNR. Compressed sensing was demonstrated by Lustig et al. in 2007,[16] and is an iterative image reconstruction technique that that reconstructs images from highly undersampled data, without the need to acquire complementary information, as is required with parallel imaging. An appropriate choice of k-space sampling pattern

prevents loss of SNR, but image reconstruction is relatively slow and imposes a high computational burden, and it is these factors that slowed the uptake of compressed sensing in clinical environments.

### Machine Learning Techniques

The introduction of machine learning (ML) techniques has dramatically altered the joint task of MR data acquisition and image reconstruction, and is a rapidly growing field. In general, ML refers to the ability of a predictive model to learn to perform a task, and eventually make predictions based on unseen input data. Although the up front training of an ML model can be time and resource expensive, the resulting pipeline can reconstruct images significantly faster than other reconstruction techniques (e.g. compressed sensing). As such, numerous ML models have been developed to perform tasks that mitigate the issues that arise from undersampling k-space. Some approaches include estimating the missing k-space data,[17] denoising the resulting images,[18,19] or learning the transformation from data space to artifact-free image space.[20]

Super resolution (SR) techniques are an ML method that originated in the field of computer vision, and generate high resolution (HR) images from low resolution (LR) ones. Over time, numerous model architectures have been used to perform the SR task, starting from the first application of a convolutional neural network (CNN) to SR (i.e. SRCNN) in 2014,[21] to the use of a residual neural network (ResNet),[22,23] to a generative adversarial network (GAN), termed SRGAN.[24] Further description of these architectures is provided in Chapter 2.3.

In recent years, existing deep-learning super resolution models have been applied, and new ones developed, for application to LR MR images. While SRCNN was originally designed to operate on 2D natural images, Pham et al. designed a pipeline to apply SRCNN to 3D MR brain images, enabling doubling of the acquisition voxel size, and finding improved results compared to cubic spline interpolation.[25] Chen, Xie et al. developed a 3D model based on a dense ResNet for 3D MRI SR,[26] and subsequently introduced a discriminator network to their model, producing a 3D GAN that creates 3D MRI volumes nearly indistinguishable from the HR target.[27] More

recently, SR methods have been applied to clinical systems; Siemens Healthcare has released Deep Resolve Sharp,[28] a CNN that enables an increase of acquired voxel size of up to a factor of two along both in-plane axes, thereby reducing scan times.

As demonstrated thus far, many recent efforts in the areas of data acquisition and reconstruction focus on reducing scan time without degrading images significantly. In addition to improving patient experience, a decrease in scan time reduces image artifacts resulting from patient motion that may require repeated acquisitions. Furthermore, scan time reduction has the important benefits of increased patient throughput and more efficient use of costly healthcare resources.

Overall, each of the developments outlined in Section 1.2 have provided more opportunities to bridge the gaps in the trade-off between SNR, resolution, and scan time (Figure 1.1). Though previously overlooked, these advancements are now being applied to low-field systems, creating a resurgence in research and development of low-field MRI.

## 1.3 Advantages and Disadvantages of Low-Field MRI

The range of field strengths used on clinical systems today is accompanied by a corresponding continuum of advantages and disadvantages. In the low field regime ($B_0 < 1.5$ T), industry has found a "sweet spot" of operation at $B_0 \approx 0.5$ T where there are opportunities to leverage the benefits of low field while still balancing the need for sufficient raw signal strength.[7,29] What follows is an overview of some of the advantages and disadvantages of utilizing MR systems at $B_0 \approx 0.5$T with high-performance components in comparison to conventional (1.5T and 3T) systems.

### 1.3.1 Installation and Operation

The installation and implementation of conventional MR systems present a number of challenges. MRIs can weigh between ≈4 000 to ≈40 000 kg and require specialized equipment to move them to their designated location. In addition, such systems have a large footprint and typically require the venting of cryogens, making implementation in existing imaging departments difficult. These barriers, along with the cost of purchase, installation, maintenance, and operation, hinder accessibility and

ultimately limiting patient access.[2,30]

On the other hand, the installation and infrastructure requirements for low-field systems are significantly reduced. They are generally lighter (e.g. $< 3600$ lbs[31]) and smaller, minimizing siting requirements. Additionally, many low field systems use closed, conduction cooled superconducting magnets that do not require venting of cryogens, enabling easier installation and dramatically reducing the need for scarce, expensive liquid helium coolant.[7,32] The possibility of installing such systems in existing departments opens the door to new applications of MRI outside of diagnostic imaging. Such systems could, for example, be installed in existing emergency departments or intensive care units, allowing for efficient, point-of-care imaging of patients that require time-sensitive decision making, or that are too ill to tolerate transport.

### 1.3.2 Patient Safety and Comfort

MRI is generally considered to be a safe form of imaging compared to modalities such as computed tomography (CT) and x-ray, which expose a patient to potentially harmful ionizing radiation. There are, however, safety hazards associated with the strong main magnetic field, RF field, and quickly varying gradient fields[33] used to generate an MR image.

In general, the higher field strength a magnet produces, the larger the fringe field it creates. This increases the risk of metallic objects being attracted to the magnet and becoming dangerous projectiles. A low-field magnet possesses a smaller fringe field, and the risk of patient injury due to projectiles is significantly reduced.

During imaging, multiple RF pulses are required for each image acquisition, each depositing energy into a patient's body. The resulting absorption of energy by various tissues is termed the specific absorption rate (SAR), and is estimated by all commercial MR systems to ensure patient safety during a given imaging sequence. In addition to depending on RF field characteristics, SAR is proportional to the square of the static field strength. For a given imaging sequence, SAR is therefore 36 times lower at 0.5T than at 3T, resulting in a reduced risk of heat-induced injuries at low field strengths. Similarly, imaging with lower field strength offers the advantage of reduced heating of implants, and interventional devices (e.g. catheters, guidewires[34]), as the wavelength of the RF wave required to generate a signal increases beyond the

length of these devices and beyond dimensions of the human body.

The possibility of PNS must be considered when utilizing strong gradients at any field strength. PNS can occur when rapidly changing gradient fields induce an electric field in the human body, causing stimulation of peripheral nerves. This imposes limits on the parameters of imaging sequences that require rapid gradient switching such as echo planar imaging (EPI) and steady-state free precession (SSFP). Notably, PNS concerns are significantly reduced when using systems designed exclusively for neuroimaging due to a reduction in the time-changing gradient fields at relevant anatomy.[35] Such head-only systems have the additional advantage of reducing a patient's experience of claustrophobia during imaging.

Finally, Ampére's force law describes forces imposed on the gradient coil windings during scanning, which can produce sound pressure in excess of 100dB, which comes with the risk of temporary or permanent hearing loss, or tinnitus.[33] Even when operating with high gradient strengths, these forces are reduced at low field, which thereby reduces the acoustic noise, allowing for less risk of acoustic injury and a more relaxed experience for the patient.[30,36]

### 1.3.3 Image Acquisition

The number and timing of RF pulses combined with the use of gradient fields defines the MR imaging *sequence*. Sequence parameters are constrained by physiological limits as well as what the system's components permit, including $B_0$, and each combination confers different advantages and disadvantages. The following is an overview of aspects to be considered when regarding $B_0$ as a changeable imaging parameter.

#### General Considerations

It is necessary that the RF field experienced by hydrogen nuclei be consistent across the region to be imaged for a uniform signal to be measured from a given type of tissue.

First, the intensity of an electromagnetic field decreases as it passes through the tissues of the human body. The depth at which a wave's intensity drops to 37% of its initial value is known as its penetration depth, and decreases with an increase in the frequency of the wave. Consequently, since the frequency of the RF pulse used

in MRI increases in proportion to the main magnetic field strength, the penetration depth of the RF wave decreases as the field strength increases. On its own, this effect, therefore, leads to a decrease in the signal acquired from regions deep inside the body as field strength increases. Furthermore, as the frequency of the RF wave increases, the wavelength decreases proportionally. Consequently, an increase in field strength, and therefore in the frequency of the RF pulse required, leads to the wavelength of the RF pulse becoming comparable to the dimensions of the human body; as $B_0$ increases to 3T, the wavelength decreases to the size of the human torso, and at 7T it becomes comparable to the size of the human head. In such scenarios, a standing wave can form which causes the intensity of the RF wave to increase with depth. Combined, these two effects can lead to complicated RF excitation in the region being imaged, when imaging is performed at field strengths above conventional values. Fortunately, both effects are reduced at low field strengths, and can ultimately be considered negligible. As discussed in Section 1.3.2, SAR is proportional to the square of the static field strength. Since the contribution of $B_0$ to SAR is reduced at low field strengths, there are more options for the strength and length of time of the RF pulse used, thereby permitting the use of sequence parameters that would be SAR limited at high field.

The consideration of concomitant gradient fields is of particular importance when utilizing a low-field system with strong gradients. Maxwell's equations dictate that any gradient field must be accompanied by concomitant fields with amplitudes comparable to that of the applied gradient's amplitude itself. These additional fields can, therefore, alter the overall desired magnetic field by a non-negligible amount when imaging with low $B_0$ and high amplitude gradient fields,[37] leading to distortions and signal drop out in acquired images due to errors in spatial encoding. Fortunately, methods exist to mitigate these errors, including compensatory frequency or shim offsets[38] and sequence-based corrections.

### Artifact or Imaging Opportunity?

The behaviour of spins in one field-strength regime can bestow desired results while creating undesired artifacts in another. As such, the choice of field strength must be considered based on the task at hand. Two examples relevant to a comparison of low and high field strength acquisitions are presented here.

The local magnetic field experienced by a given spin changes slightly depending on its molecular environment. Despite being placed in the same magnetic field, the hydrogen nuclei in water, for example, experience a slightly different magnetic field than hydrogen nuclei in fat, due to differences in magnetic shielding. This difference in magnetic field leads to different resonant frequencies, which itself can lead to errors in spatial encoding efforts. The magnitude of this difference is proportional to $B_0$, therefore artifacts resulting from signal mismapping are easier to prevent at low field strengths. Some imaging techniques, however, rely on this difference to be successful. Fat suppression, for example, is necessary for tumour evaluation in the high-fat region of the optic nerve; imaging sequences that perform fat suppression use the difference in spin behaviour to null any signal emanating from hydrogen nuclei in fat, while still measuring the signal in other tissues.

A tissue with high magnetic susceptibility will interact with $B_0$ and therefore distort the local magnetic field. In the context of anatomical MR imaging, susceptibility differences between tissues can cause spins to behave in unsought ways, causing errors in spatial encoding and therefore creating signal dropout and spatial mismapping in the images acquired. Artifacts of this nature are especially prevalent around implants and air/tissue boundaries, but the magnitude of this effect is proportional to $B_0$, and is therefore reduced at low field strengths. However, the effects of magnetic susceptibility can also be leveraged for other acquisition strategies: functional MRI, for example, relies on the susceptibility difference between oxygenated and deoxygenated blood in and around blood vessels and achieves best results at high field strengths.

### Image Contrast

Though a small effect, the behaviour of spins at low field strength can alter the desired image contrast. Following an RF pulse, the recovery of spins to the longitudinal plane is governed by the $T_1$ time constant, with different tissues exhibiting different $T_1$ values, and imaging parameters can be chosen to probe this difference. It has been shown that $T_1$ values typically decrease at low field strengths,[34,39] which could permit a reduction in overall exam time.[29] On the other hand, decreases in T1 values results in smaller absolute differences between tissues, and poses a challenge for imaging sequences that rely on these differences to produce contrast (e.g. T1 weighted

and T1 fluid attenuated inversion recovery sequences). Further, contrast-enhanced imaging relies on the properties of gadolinium-based contrast agents to reduce the $T_1$ of tissues in which it is taken up. A reduction in the $T_1$ of surrounding tissues means that a further reduction is more difficult to detect, and contrast enhancement is less pronounced at low field strengths.

### Scan Times

Despite its flexibility and superior soft-tissue contrast, long scan times are an inherent aspect of MR imaging that puts it at a disadvantage to other imaging modalities. As outlined in Sections 1.3.1 and 1.3.2, low-field systems have the potential for use in emergency and point-of-care scenarios. However, long scan times may be a larger issue in this field-strength regime, where the reduced SNR often leads to the need for multiple measurements and signal averaging. Addressing this concern may therefore be necessary to increase patient comfort and reduce the possibility of motion artifacts, and to bolster the value of low-field MRI in an emergency or point-of-care scenario.

## 1.4 Balanced Steady-State Free Precession Imaging at Low Field Strengths

Balanced Steady-State Free Precession (bSSFP) is an imaging technique that was proposed by Oppelt et al. in 1986,[40] but didn't find application until fast, strong, and precise gradient systems were developed. It has recently seen a resurgence at low field strengths that utilize high-performance gradients[41] as it is distinctly suited to demonstrating the advantages of imaging with such systems.

Importantly, bSSFP sequences measure a coherent steady-state signal (see Section 2.2.3), and therefore offer the most SNR per unit time of all imaging sequences.[42,43] Use at low field strengths then ensures that the raw signal available is used to its full potential while still enabling high resolution imaging in clinically acceptable exam times. Further, the signal is highly dependent on the flip angle used. The reduced SAR at low field strengths means that the duration for which the RF pulse is transmitted can be increased, permitting a wider range of flip angles without reaching patient heating limits and allowing for sequence parameters that optimize the available signal. In addition, the signal measured in bSSFP acquisitions inversely depends

on T1 values, therefore with all other sequence parameters constant, reduced T1 at lower fields leads to a higher amplitude of transverse magnetization available for measurement.

Notably, bSSFP sequences are particularly sensitive to inhomogeneities in $B_0$ that can be caused by susceptibility differences at tissue interfaces. bSSFP images acquired under these conditions demonstrate band-like artifacts that can obscure relevant anatomy. The nature of the bSSFP sequence dictates that banding occurs when spins are off-resonance by frequencies of $\pm 1/2TR$. While methods such as phase cycling of the RF pulse[44] (Section 2.2.3) can be used to compensate for the effects that cause banding, they require longer scan times, leading to motion-induced artifacts, as well as increased reconstruction times. bSSFP sequences can therefore benefit from the increased $B_0$ homogeneity and decrease in susceptibility effects at low field strengths.[41] Performing bSSFP acquisitions on head only systems with high strength gradients has the additional advantage of allowing for a reduction in TR, which can limit the time available for off-resonance precession to occur, thereby reducing banding artifacts.

Two regions that could benefit from the advantages of low-field bSSFP imaging are the temporal bone and the paranasal sinuses. These regions are typically considered challenging to image with MRI due to many air/bone/tissue interfaces that generate off-resonance effects due to magnetic susceptibility differences. In addition, these areas are comprised of complex anatomical structures that require high contrast and resolution to visualize effectively.

## 1.5 Imaging the Temporal Bone and Paranasal Sinuses

### 1.5.1 Temporal Bone

The temporal bones are two major bones of the skull that form its lateral base and protect the structures and nerves of the ear canal, middle ear, and inner ear. The vestibulocochlear and facial nerves begin at the cerebellopontine angle in the brain before passing through the internal auditory canal (IAC) to enter the inner ear. The vestibulocochlear nerve splits into the vestibular and cochlear nerves at the distal end of the IAC before abutting the cochlea and vestibule, respectively. Together with the

semicircular canals, the cochlea and vestibule make up the bony labyrinth of the inner ear (Figure 1.3). Inside the bony labyrinth is a collection of fluid-filled chambers and tubes called the membranous labyrinth, responsible for hearing and balance. The cochlea converts sound waves into nerve signals, while the vestibule and semicircular canals are responsible for the sense of balance and equilibrium.[45] After entering the inner ear, the facial nerve winds it way between the vestibule and cochlea before exiting the inner ear through the stylomastoid foramen to the facial canal.

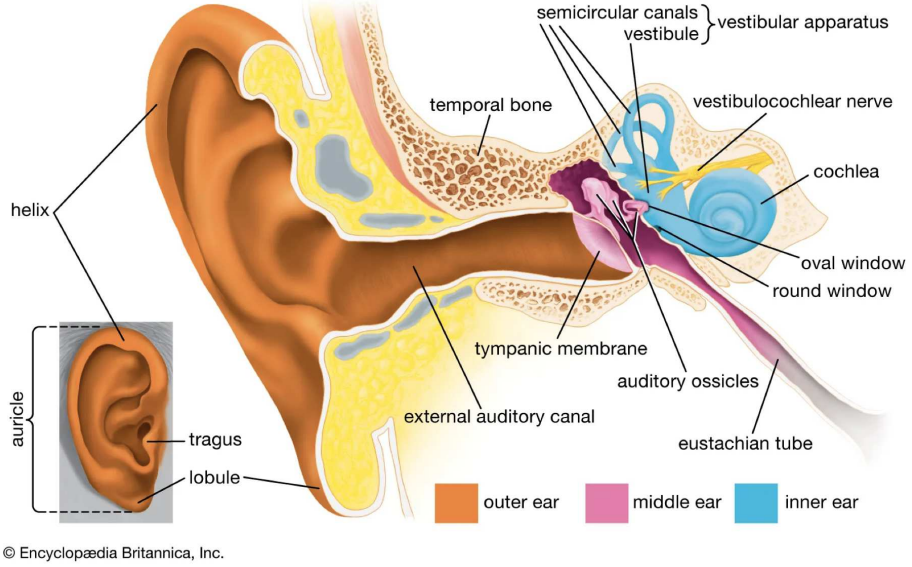


Figure 1.3: The temporal bone protects the structures of the inner ear, depicted here alongside the middle and outer ear. The figure is reproduced from Encyclopedia Britannica.[46]

This complex space is imaged for a variety of reasons including evaluation of sensorineural hearing loss, evaluation of cochlear implant candidacy, and monitoring neoplastic and infectious processes. Patients with conditions of the inner ear may present with a myriad of symptoms including vertigo, autophony, tinnitus, or decreased speech discrimination,[47] and often both CT and MRI are ordered as they tend to provide complimentary information. CT, for example, has become the gold standard for diagnosis of dehiscence of the superior semicircular canal,[48] while the superior soft tissue contrast and isotropic resolution of MRI aids in detection of small schwannomas along the IAC.[49]

The contrast and resolution provided by bSSFP sequences are ideal for visualizing the complex region of the inner ear, as the signal from the fluid-filled structures is

high compared to the petrous part of the temporal bone. However, the potential formation of banding artifacts traversing the relevant structures leads to preference for lower SNR sequences, such as fast spin echo (FSE), when imaging at conventional field strengths; Benson et al. utilize an FSE sequence to image the inner ear and note that banding artifacts are particularly problematic when examining the vestibule, semicircular canals, and cochlea with bSSFP at 3T.[50] Similarly, a study performed at 1.5T involving radiologists' ratings of "detectability of inner ear structures" showed preference for FSE acquisitions, and the authors hypothesize that the banding artifacts present in bSSFP acquisitions were to blame for their inferior ratings.[51] With increased field homogeneity and reduced susceptibility effects, low-field bSSFP imaging of the temporal bone has the potential to provide high resolution, artifact tolerant results.

### 1.5.2 Paranasal Sinuses

The paranasal sinuses are four sets of air-filled spaces that extend from the nasal cavity, each named after the bone that they are situated: the frontal, ethmoid, maxillary, and sphenoid sinuses (see Figure 1.4). The role of the paranasal sinuses is not definitively known, however they are thought to increase the resonance of the voice, heat and humidify air that is breathed in, and protect vital structures in the case of trauma.

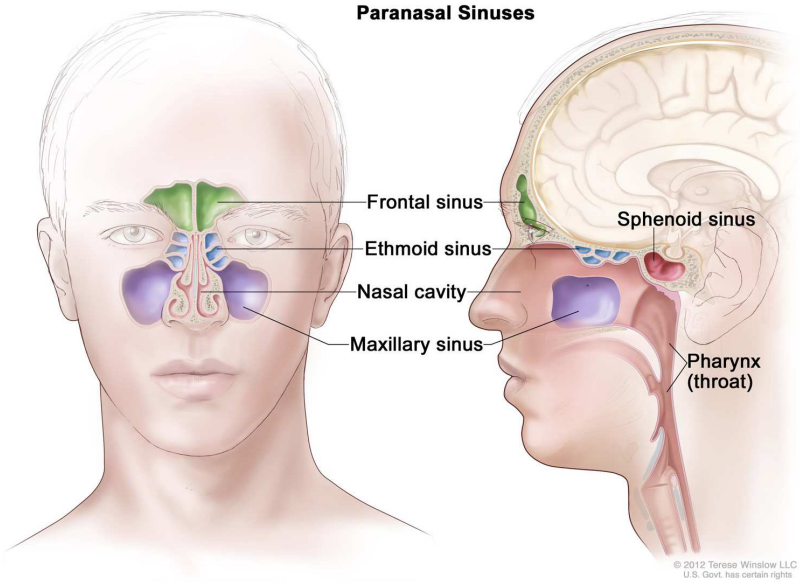


Figure 1.4: The locations of the nasal cavity and four sets of paranasal sinuses. Figure is reproduced from the National Institute of Health's National Cancer Institute.[52]

The nasal cavity and sinuses, as well as the drainage pathways connecting them, are lined with a mucous membrane known as mucosa, responsible for keeping the sinuses moist and for trapping airborne pathogens. Inflammation of this lining is termed sinusitis, and is a common condition encountered by primary care physicians.[53] Patients experiencing acute or chronic sinusitis may present with sinus pain and tenderness, nasal discharge or congestion, and fever.

Imaging of the paranasal sinuses and surrounding structures is indicated for patients with symptoms of sinusitis that do not respond to medical treatment, and is required for diagnosis, assessment of anatomic variants, and to provide detailed anatomy for surgical planning.[54,55] When endoscopic surgery is required, pre-surgical visualization of the cribriform plate is of particular importance. The cribriform plate is a thin portion of the ethmoid bone that forms the roof of the nasal cavity, separating it from the brain. Puncture of the cribriform plate can lead to cerebral spinal fluid (CSF) leak, olfactory impairment, and frontal lobe injury.[56]

CT is the modality of choice for imaging of the paranasal sinuses, due primarily to its availability, speed, and the resolution of bony structures that is provides. CT, however, comes with significant disadvantages including exposure of the patient to

potentially harmful ionizing radiation and possible reduced resolution in the superior-inferior direction. Access to MR imaging could prevent exposure of the patient to radiation, and a bSSFP acquisition performed on a low-field system could provide the SNR efficiency, isotropic resolution, artifact tolerance, and speed required to compete with CT for imaging the paranasal sinuses.

## 1.6 Research Objectives

As described, low-field MRI has the potential to provide new opportunities to image challenging regions which demonstrate susceptibility-induced off-resonance effects. As such, this thesis investigates the application of bSSFP sequences to imaging of the temporal bone and paranasal sinuses on a 0.5T, head-only system with high strength gradients. It is comprised of three studies, each addressing a unique research objective.

The first study is presented in Chapter 3 and addresses the first objective - to quantitatively evaluate the images resulting from 3D bSSFP protocols designed to image the temporal bone and paranasal sinuses at 0.5T with clinical resolutions. More specifically, the artifact tolerance and the SNR achieved will be quantified. The study is titled

**Artifact-resistant balanced steady-state free precession imaging of the temporal bone and paranasal sinuses without phase-cycling at 0.5T**

While the metrics examined in the first study are informative, the true value of an MR acquisition lies in a radiologist's ability to use the image to answer a clinical question. This aspect is investigated in the second study, presented in Chapter 4, where the objective is to evaluate radiologists' ability to visualize structures of the inner ear in bSSFP images acquired at 0.5T in comparison to clinical bSSFP acquisitions at 3T. The second study is titled

**Low-field vs. conventional field balanced steady-state free precession imaging of the temporal bone: radiologist rating of anatomical visualization.**

The third study is presented in Chapter 5 and its objective is to explore the benefits of low resolution bSSFP acquisitions, and to return the resulting images

to high resolution, without significant degradation, with a publicly available super resolution model. The third study is titled

**Super resolution allows for the benefits of low resolution balanced steady-state free precession imaging without degrading image quality.**

Independent hypotheses for each study are provided in their corresponding chapters.

# Chapter 2

# Theory and Experimental Details

## 2.1 Acquiring an MR Image

### 2.1.1 Signal Generation

Hydrogen nuclei are found in abundance in fat and water in the tissues of the human body. The hydrogen nucleus possesses a fundamental property known as "spin", or spin angular momentum, and for this reason hydrogen nuclei are often referred to as "spins" in MR literature. The angular momentum of a spin generates a magnetic dipole moment, and it is this property that is used to generate MR images. Without any external influence, the direction of a given magnetic moment is random. However, the presence of a strong, external magnetic field of strength $B_0$, causes a small excess of spins to become aligned along its direction. The direction of the main magnetic field in MR imaging is conventionally parallel to the the $z$-axis, i.e. $\vec{B}_0 = B_0\hat{z}$, where $\hat{z}$ is a unit vector along the $z$-axis. The excess of dipoles aligned along $\hat{z}$ results in a net magnetization vector, $\vec{M}$, the magnitude of which depends on many factors including the density of hydrogen nuclei, and the strength of $B_0$. Generally speaking, any external magnetic field, $\vec{B}$, will induce a torque on $\vec{M}$, causing it to precess about the direction of the field with a frequency known as the Larmor frequency, $\omega = \gamma\vec{B}$, where $\gamma$ is the gyromagnetic ratio ($\gamma = 42.6$ MHz/T for hydrogen nuclei). Spins oscillating at this frequency are said to be *on resonance*, and this motion is described by the 3D equations of motion for $M$ known as the Bloch equations:

$$\frac{d\vec{M}}{dt} = \vec{M} \times \gamma\vec{B}. \tag{2.1}$$

To generate a measurable signal, a secondary oscillating magnetic field is applied for a short duration. This field is set to oscillate at the Larmor frequency, which is in the RF range for conventional clinical field strengths. The pulsed magnetic field

is therefore referred to as an RF pulse. The application of the RF pulse tips the net magnetization away from $\hat{z}$ such that a transverse component, $M_{xy}$, is produced. Since $M_{xy}$ continues to oscillate about $\hat{z}$, it is convenient to consider the transverse plane as complex such that

$$M_{xy} = M_x + iM_y. \tag{2.2}$$

Immediately following the RF pulse, $\vec{M}$ will relax back to the longitudinal plane as more and more spins return to their equilibrium positions aligned with $B_0$. The longitudinal component, $M_z$ recovers to its equilibrium value, now referred to as $M_0$, according to a relaxation time constant, $T_1$. Simultaneously, but independently, $M_{xy}$ decays due to atomic and molecular interactions creating local disturbances in the magnetic field. The variation in field strength experienced by individual spins causes the spins that make up $M_{xy}$ to start oscillating out of phase. This process is governed by the $T_2$ time constant. However, other field disturbances can exist due to inhomogeneities in the main magnetic field, which accelerate the dephasing, and therefore the decay, of the transverse magnetization. This contribution to decay is governed by the $T_2^{'}$ time constant. Combined, these two effects are described by $T_2^*$, where

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2^{'}} \tag{2.3}$$

Notably, $T_2^*$ is always less than $T_1$. Differences in $T_1$, $T_2$, $T_2^*$, and spin density (or a combination of these) between tissues can be probed with various imaging strategies, and it is these differences that generate contrast in MR images.

Equation 2.1 must now be modified to describe the precession of $\vec{M}$, as well as its relaxation to its equilibrium position. For simplicity, we will assume that the effects of field inhomogeneities are negligible such that $T_2^* = T_2$:

$$\frac{d\vec{M}}{dt} = \vec{M} \times \gamma\vec{B} + \frac{(M_0 - M_z)\hat{z}}{T_1} - \frac{(M_x\hat{x} + M_y\hat{y})}{T_2} \tag{2.4}$$

where $\hat{x}$, $\hat{y}$, and $\hat{z}$ are unit vectors along the $x$, $y$, and $z$ directions. Equation 2.2 can then be used to solve the first order linear differential equations defined by

Equation 2.4 to obtain

$$M_{xy}(\vec{r},t) = M_{xy}(\vec{r},0)\exp\left[\frac{-t}{T_2}\right]\exp[-i\omega(\vec{r})t] \tag{2.5}$$

and

$$M_z(\vec{r},t) = M_0(\vec{r}) + [M_z(\vec{r},0) - M_0(\vec{r})]\exp\left[\frac{-t}{T_1}\right], \tag{2.6}$$

where $M_{xy}(\vec{r},0)$ is the transverse magnetization at location $\vec{r}$ immediately following an RF pulse, and is equal to $M_0(\vec{r})\sin\alpha$, where $\alpha$ is the angle that $\vec{M}$ is tipped away from the $z$-axis. $M_z(\vec{r},0)$ is therefore given by $M_z(\vec{r},0) = M_0(\vec{r})\cos\alpha$.

Importantly, it is this oscillation of $M_{xy}(\vec{r},t)$ that is fundamental in generating an MR image; the oscillating magnetic field in the transverse plane generates an electric current in the RF receive coils, as governed by Faraday's law of electromagnetic induction.

### 2.1.2 Data Acquisition and Image Reconstruction

Clearly, spatial information must be encoded into the spins so that the location of a given signal can be transferred to the image. To do this, linearly varying gradient fields are applied in the $x-$, $y-$, and $z-$directions so as to modify $B_0\hat{z}$ in a specified manner. The magnetic field experienced by a spin at a location $\vec{r}$ and time $t$ is then given by

$$\vec{B}(\vec{r},t) = [B_0 + G_x x + G_y y + G_z z]\hat{z} = B_0 + \vec{G}\cdot\vec{r} \tag{2.7}$$

for a time-independent gradient field of magnitude $G$. The addition of gradient fields modifies the Larmor frequency accordingly:

$$\omega(\vec{r}) = \gamma(B_0 + \vec{r}\cdot\vec{G}) \quad = \omega_0(\vec{r}) + \gamma\vec{r}\cdot\vec{G} \tag{2.8}$$

For convenience, we can now adopt a reference frame that rotates with the net magnetic field vector about the $z-$axis with frequency $\omega_0$. In this frame of reference

$$\omega(\vec{r}) = \gamma\vec{r}\cdot\vec{G} \tag{2.9}$$

As demonstrated, the application of gradient fields causes spins at different locations, $\vec{r}$, to precess at different frequencies. Consequently, application of a gradient field will cause spins to accrue various phase offsets, $\phi$, according to their local magnetic field and the time over which the field is experienced. The phase offset is therefore

$$\phi(\vec{r},t) = \int_0^t \omega(\vec{r},t')dt' = \gamma \int_0^t \vec{r}\cdot\vec{G}(t')dt' = \gamma t\vec{r}\cdot\vec{G} \tag{2.10}$$

for time-independent gradient fields.

The signal measured at time $t$ is the vector sum of signals generated from transverse magnetization, at all points in the volume being imaged. Combining equations 2.5 and 2.10, we can therefore state that the signal is given by

$$s(t) = \int M_{xy}(\vec{r},0)\exp\left[\frac{-t}{T_2}\right]\exp[-i\phi(\vec{r},t)]d\vec{r} \tag{2.11}$$

This equation can be simplified if the time between signal generation (via the RF pulse) and measurement time, $t$, is significantly less than $T_2$, such that $\exp[-t/T_2]$ approaches 1. Further, as we know that $M_{xy}(\vec{r},0)$ results from the RF tipping of $M_0$, we know that it is proportional to the density of spins, temperature of the region, and the static magnetic field $B_0$, as well as RF receive coil sensitivity. We can therefore define an "effective spin density" term, $\rho(\vec{r})$ such that :

$$s(t) = \int \rho(\vec{r})\exp[-i\phi(\vec{r},t)]d\vec{r}. \tag{2.12}$$

Now, if we define

$$\vec{k}(t) = \frac{\gamma}{2\pi}\int_0^t \vec{G}(t')dt', \tag{2.13}$$

we can re-state $\phi(\vec{r},t)$ (Equation 2.10) as

$$\phi(\vec{r},t) = \gamma\int_0^t \vec{r}\cdot\vec{G}(t')dt' = 2\pi\vec{r}\cdot\vec{k}(t), \tag{2.14}$$

leading to

$$s(\vec{k}) = \int \rho(\vec{r})\exp[-i2\pi\vec{r}\cdot\vec{k}(t)]d\vec{r}. \tag{2.15}$$

This result demonstrates that the acquired signal, $s(\vec{k})$, is the Fourier transform of the effective spin density, $\rho(\vec{r})$, which is what we aim to visualize. The inverse Fourier transform therefore allows us to obtain an image,

$$\rho(\vec{r}) = \int s(\vec{k})\exp[i2\pi\vec{r}\cdot\vec{k}(t)]d\vec{k}. \tag{2.16}$$

This equation therefore returns the relative "amount" that a spatial frequency vector, $\vec{k}(t)$ contributes across space $\vec{r}$. Hence, the data collected in MRI is a series of complex numbers, placed in a data array known as *k-space*. The k-space matrix has axes $x$, $y$, and $z$ which intersect at the center of k-space where $x = y = z = 0$.

### 2.1.3 Discrete Sampling of k-Space

Importantly, we have thus far considered $\vec{r}$ and $\vec{k}$ as continuous variables, leading us to Equations 2.15 and 2.16. However, truly continuous sampling of $\vec{k}$ is not possible in practice, and as such, k-space is a discrete, finite data array. Assuming we acquire $N_x$, $N_y$, and $N_z$ points in k-space in the x-, y-, and z- dimensions, we can discretize Equations 2.15 and 2.16 to find

$$s(k_x, k_y, k_z) = \sum_{x=0}^{N_x-1}\sum_{y=0}^{N_y-1}\sum_{z=0}^{N_z-1}\rho(x,y,z)\exp\left[-i2\pi\left(\frac{k_x x}{N_x}+\frac{k_y y}{N_y}+\frac{k_z z}{N_z}\right)\right] \tag{2.17}$$

$$\rho(x,y,z) = \frac{1}{N_z N_y N_z}\sum_{x=0}^{N_x-1}\sum_{y=0}^{N_y-1}\sum_{z=0}^{N_z-1}s(k_x,k_y,k_z)\exp\left[i2\pi\left(\frac{k_x x}{N_x}+\frac{k_y y}{N_y}+\frac{k_z z}{N_z}\right)\right] \tag{2.18}$$

where the factor of $1/N_x N_y N_z$ in Equation 2.18 is required to ensure the total power in image space is equivalent to the total power in the signal (Parseval's theorem).

As demonstrated by Equation 2.13, different values of $k$ are defined by the integral of the function describing gradient strength over time. It is therefore with careful manipulation of $\vec{G}$ that k-space gets filled. If, for example, a constant gradient is applied to alter $B_0$ along the x-axis, and measurements are made with time interval $\Delta t$, then a row of k-space can be filled according to $k_{nx} = G_x n_x \Delta t$, where $n_x =$

$(0, 1, ..., N_x/2)$. This gradient application has the effect of systematically altering the frequency of spins, and as such this process is termed *frequency encoding*. In practice, a negative amplitude gradient is often applied prior to data acquisition, moving to the edge of k-space and allowing for an entire row of data to be acquired in succession, i.e. $k_{nx} = -G\frac{N_x}{2}\Delta t + G_x n\Delta t$, where $n = (0, 1, ..., N_x)$. To change the y- position in k-space, a second gradient is applied for a fixed duration, $t$, according to $k_{ny} = n_y \Delta G_y t$. This gradient application has the effect of systematically changing the phase of spins along the y-direction, and is therefore termed *phase encoding*. Following the application of the phase-encode gradient, frequency encoding can be repeated to collect data for the corresponding row of k-space. An example of basic frequency and phase encoding for a 2D image is shown in a pulse sequence diagram in Figure 2.1. Pulse sequence diagrams are read from left to right with each row depicting the amplitude and length of time with which a given gradient is applied.

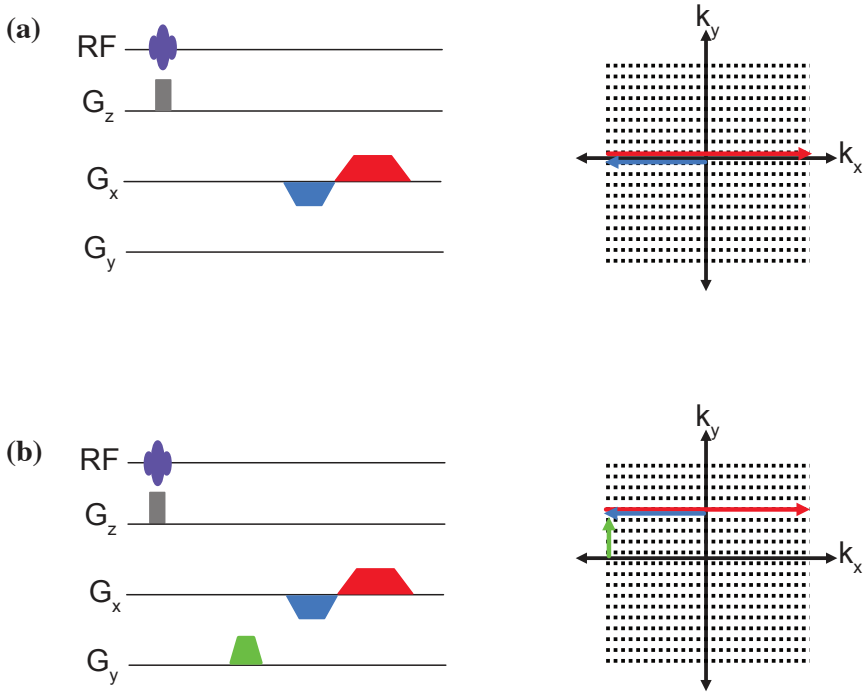


Figure 2.1: (a) and (b) both show the basic sequence of events required to traverse a desired row of k-space. In both, the application of a slice-select gradient, $G_z$, occurs simultaneously with an RF pulse designed to excite spins within the desired slice. In (a), a negative frequency gradient, $G_x$, moves the k-space point being measured from the origin to the negative periphery. A subsequent positive application of $G_x$, with twice the duration of the previous pulse, allows for traversal of the entire row of k-space. In (b), a phase-encoding gradient, $G_y$, is applied to move along the positive y-axis of k-space. The frequency gradient is then applied as it was in (a) to traverse the row.

If a 3D acquisition is desired (as described by Equation 2.18), further spatial encoding can be accomplished by additional phase encoding along the third dimension.

The scaling property of the Fourier transform defines the image field of view (FOV) in each dimension to be equivalent to the reciprocal of the step size in k-space taken in that dimension. For acquisition of a 2D image slice, we therefore have the relationships

$$FOV_x = \frac{1}{\Delta k_x}, \quad FOV_y = \frac{1}{\Delta k_y}, \tag{2.19}$$

which show that the size of $\Delta k$ is limited to the reciprocal of the size of the object

in the corresponding dimension. If $\Delta k$ is too large, the FOV may not encompass the anatomy to be imaged, causing overlapping of signal, or aliasing.

We can also express the FOV of an image via the intuitive relationship between the dimensions of an image's pixels and the number of pixels acquired i.e. $FOV_x = N_x \Delta x$ and $FOV_y = N_y \Delta y$. This leads us to expressions for the *resolution* of an image, which is:

$$\Delta x = \frac{FOV_x}{N_x}, \quad \Delta y = \frac{FOV_y}{N_y}. \tag{2.20}$$

From this relationship, we can see that the higher the number of data points acquired, the smaller the voxel size, or the higher the image resolution. This is equivalent to increasing the extent of k-space measurements made. The outer edges of k-space therefore contain information on the high frequency components of on an image, and those at the center contain the information related to low frequency information, or contrast. This concept is demonstrated in Figure 2.2.

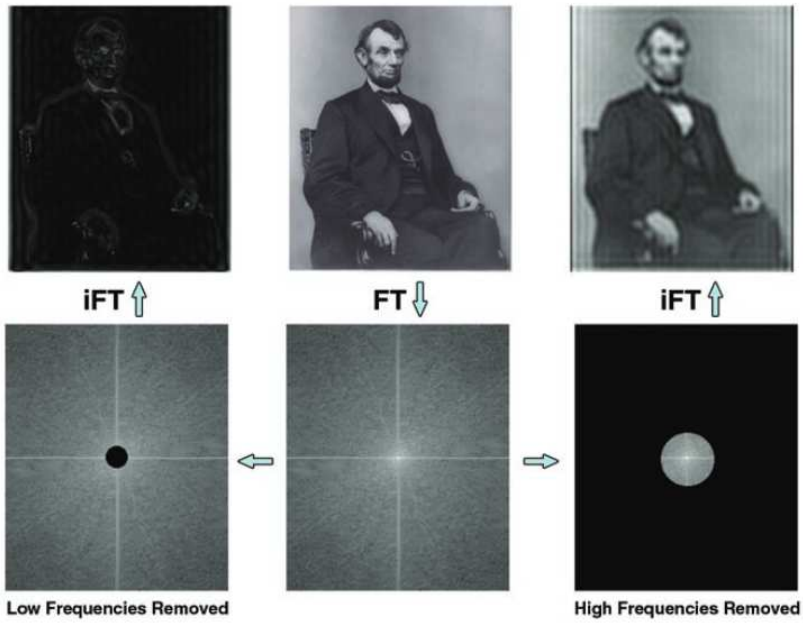


Figure 2.2: K-space data (lower row) and the corresponding image (upper row) after inverse Fourier transformation. With low frequencies removed from the center of k-space (left-most column), the image contains only high frequency information such as edges. When only the central region of k-space is sampled (right-most column), the high frequency information is lost, resulting in an image with sufficient contrast, but low image resolution. Figure is reproduced from Gallagher et al. 2008.[57]

High resolution acquisitions can be time-consuming due to the numerous phase-encoding steps that must be taken to fill k-space. This can be mitigated, to some extent, by decreasing the number of values acquired in the periphery and instead performing zero-padding of the k-space matrix. This has the effect of reducing $\Delta x$ and $\Delta y$ as shown in Equation 2.20, thereby improving image visualization. It is important to note, however, that no new spatial information is actually acquired, and only the apparent resolution of the image is improved. Clearly, this technique has its limits, and reducing resolution may lead to loss of information that could be important for clinical diagnosis.

Finally, the discrete sampling of k-space defined by Equation 2.18 consists of a finite number of samples, $N_x, N_y$ and $N_z$. The inability to measure an infinite number of k-space points is equivalent to truncation of the "true" k-space representation of an object. In image space, this leads to an artifact known as Gibbs ringing, which appears as faint lines parallel to high contrast interfaces. Fortunately, filtering of k-space can be used to mitigate the worst of these effects.

## 2.2 Balanced Steady-State Free Precession

### 2.2.1 The Gradient Echo Sequence

The transverse magnetization described by Equation 2.5 decays exponentially according to $T_2$ in a process known as *free induction decay* (FID), therefore, multiple RF pulses are applied, separated by the *repetition time*, TR, to generate the signals required to fill k-space. When imaging with a *gradient echo* (GRE) sequence, the frequency (or "readout") gradient is applied twice with equivalent amplitudes, in opposite directions, so as to dephase and re-phase the transverse magnetization. Doing so generates a spatially encoded echo of the initial FID between RF pulses at the *echo time*, TE. A basic pulse sequence diagram, provided in Figure 2.3, shows the sequence of events used to generate a GRE.

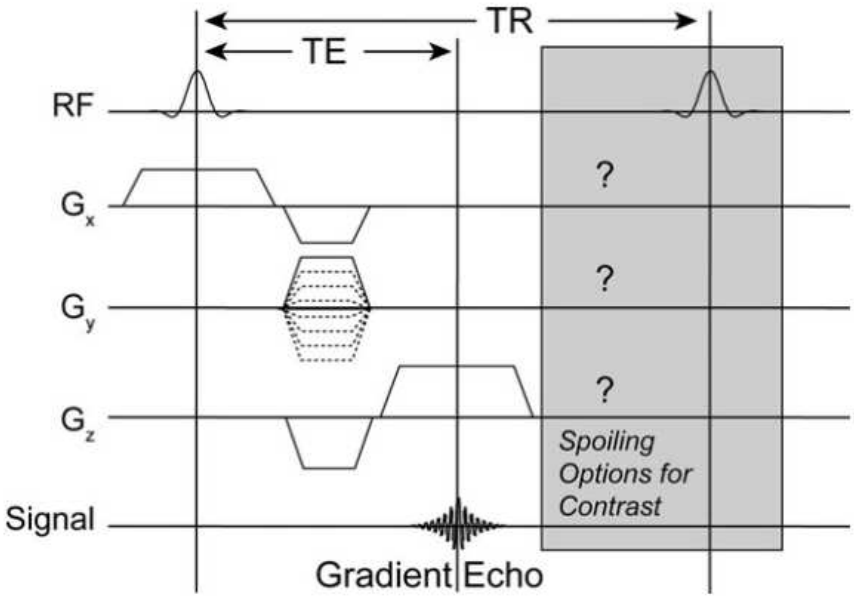


Figure 2.3: A pulse sequence diagram depicting the sequence of events that generate a gradient echo. To start, the RF pulse is applied to flip the spins simultaneously excited by the slice-select gradient, $G_x$ by an angle $\alpha$, inducing free induction decay. A negative slice-select gradient is applied to undo phase offsets introduced by slice selection, while the phase-encode gradient, $G_y$ generates the intended phase offsets for a given TR (with various amplitudes for each TR indicated by the dotted lobes). A negative frequency-encoding gradient, $G_z$, is applied to dephase the FID signal, and its amplitude is then reversed so as to read out a frequency-encoded gradient echo signal at the echo time, TE. What happens following the measurement of the echo depends on the type of image contrast desired - any residual transverse magnetization can be destroyed, or "spoiled", or it can be maintained as is the case in bSSFP. Figure is reproduced from Hargreaves 2012.[58]

### 2.2.2 Steady State Imaging

GRE sequences can enable fast imaging by the use of flip angles, $\alpha$, of less than 90°, wherein a significant proportion of longitudinal magnetization remains along the z-axis. Consequently, a subsequent RF pulse can be applied with a short TR, i.e. $TR < T_1$, while still having sufficient longitudinal magnetization to act on. When TR is such that incomplete recovery of the longitudinal magnetization occurs, a *steady state* of magnetization is formed. Most rapid GRE sequences utilize steady-states in which both incomplete $T_1$ and $T_2$ relaxation occurs,[58] thereby reducing imaging time and generating a signal that is a complicated combination of both $T_1$ and $T_2$ effects.

While numerous methods exist to create and maintain a steady-state signal,[59] the simplest is that in which an initial preparation RF pulse is applied to achieve a flip angle of $\alpha/2$. A subsequent $-\alpha$ pulse after time TR/2 is then followed by alternating $\pm\alpha$ pulses separated by TR. $T_1$ and $T_2$ relaxation effects cause a smooth dampening of the transverse magnetization over successive RF pulses, causing it to evolve towards a consistent steady-state. Once in the steady-state, the signal is maintained between $\pm\alpha$ RF pulses (Figure 2.4).

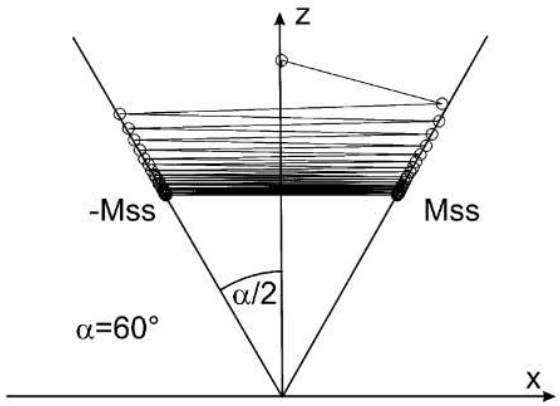


Figure 2.4: An initial $\alpha/2$ preparation pulse brings the tip of the transverse magnetization (indicated by the unfilled circles) into the $xy$-plane. Subsequent alternating $\pm\alpha$ pulses, combined with $T_1$ and $T_2$ relaxation effects cause a smooth dampening of the transverse magnetization towards an oscillating steady state. The resulting steady-state transverse magnetization is shown here to have a magnitude of $M_{ss}$. Figure is reproduced from Scheffler and Lehnhardt 2003.[42]

As described, a steady-state signal is the result of $T_1$ and $T_2$ effects, however the contrast achieved during imaging depends on how the residual transverse magnetization is managed prior to a subsequent RF pulse.

The application of a spoiler gradient induces a linear variation of spin precession across a voxel. This has the effect of reducing the magnitude of the signal, but it prevents the characteristic banding artifacts exhibited by bSSFP sequences (Section 1.4). Gradient spoiling can be used to eliminate transverse magnetization within one TR, but the dephased signal may be rephased by subsequent spoiler gradients, leading to a transverse signal that is not completely eliminated. Gradient spoiled SSFP sequences, therefore, exhibit contrast that is the result of both $T_1$ and $T_2^*$ effects.

In addition to the use of spoiler gradients, incrementally varying the phase of the RF pulse can be used to eliminate the effect of any remaining transverse magnetization. The resulting distribution of magnetization within a voxel is complicated, but the signal generated is well approximated by neglecting the residual transverse magnetization prior to RF pulses.[58] The contrast provided by sequences with RF spoiling is therefore $T_1$ weighted.

### 2.2.3 Balanced Steady-State Free Precession Imaging

When imaging with a bSSFP sequence, the residual transverse magnetization that persists is recovered before the the next RF pulse by the use of fully balanced gradients; the bSSFP pulse sequence diagram in Figure 2.5 shows that each positive gradient area is balanced by an equivalent negative gradient area, indicating that the total gradient-induced dephasing within TR is exactly zero. In other words, the steady-state magnetization vector is fully recovered prior to the subsequent RF pulse.

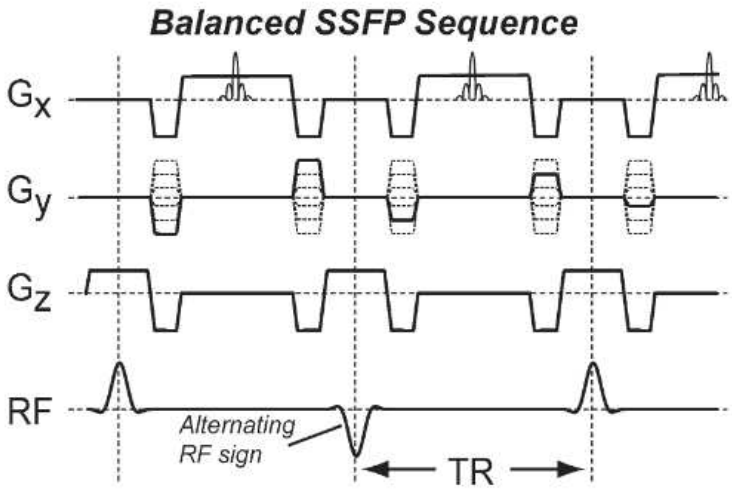


Figure 2.5: A pulse sequence diagram for bSSFP image acquisition. The net gradient amplitude between each TR is zero, such that the steady-state magnetization vector is recovered, and the effect of gradients can be ignored. Figure is reproduced from Hargreaves 2012.[58]

Figure 2.5 depicts acquisition of a single, 2D imaging slice. However, volumetric (3D) bSSFP imaging is commonly performed, as it confers certain advantages over 2D. Volumetric imaging produces an increase in SNR, as the signal is acquired from the

entire volume, and it enables the acquisition of images with high isotropic resolution. Steady-state imaging, in particular, benefits from a volumetric approach in that the steady-state magnetization need only be generated once across the entire imaging volume, and the use of short TR values permits volumetric imaging with scan times appropriate for clinical use. In the case of 3D imaging, the slice select gradient and RF pulse are used to excite a thick imaging "slab" or volume. Phase encoding then occurs simultaneously in 2 dimensions (using slice select and phase-encode gradients, for example) with all possible combinations of gradient amplitudes being cycled through. The frequency encode gradient is applied as in the 2D case.

### Signal and Contrast

To derive an expression for the transverse magnetization, $M_{xy}$, in the steady-state, we make use of the fact that relaxation effects do not change the length of the magnetization vector, $M$, within one TR, i.e. $TR < T2 < T1$. We can therefore say that

$$\frac{\delta \vec{M}}{\delta t} \cdot \vec{M} = 0 \tag{2.21}$$

Combining this with the Bloch Equations defined in Equation 2.4, we find:

$$\left(M_z - \frac{M_0}{2}\right)^2 + \frac{M_x^2 + M_y^2}{T_2/T_1} = \left(\frac{M_0}{2}\right)^2 \tag{2.22}$$

or

$$\frac{\left(M_z - \frac{M_0}{2}\right)^2}{\left(\frac{M_0}{2}\right)^2} + \frac{M_x^2 + M_y^2}{\frac{T_2}{T_1}\left(\frac{M_0}{2}\right)^2} = 1 \tag{2.23}$$

which is the equation of an ellipse centered on $(0, M_0/2)$ with width $M_0\sqrt{\frac{T_2}{T_1}}$, and height $M_0$. This ellipse defines the possible locations of the steady-state magnetization vector.

### Off-resonance Effects

Before proceeding further with the derivation, off-resonance effects must be taken into account. As discussed, bSSFP sequences are particularly sensitive to off-resonance

effects, such as field inhomogeneities in the main magnetic field, or those caused by susceptibility differences between tissues. This is due the fact that they are only capable of reversing dephasing caused by the gradients themselves, therefore any additional $T_2'$ dephasing will be incorporated into the steady-state signal.

In the reference frame oscillating at the Larmor frequency, off-resonance effects cause $M_{xy}$ to precess about $z$ by an angle $\phi$ during TR, leading $M_{xy}$ to meet the edge of the ellipse at an angle greater than $\alpha$. As such, we define an "effective flip angle", $\beta$, which represents the flip angle with the effects of precession included. Figure 2.6 demonstrates the relationship between $\alpha$, $\phi$, and $\beta$, which is

$$\tan\left(\frac{\beta}{2}\right) = \frac{\tan\left(\frac{\alpha}{2}\right)}{\cos\left(\frac{\phi}{2}\right)}. \tag{2.24}$$

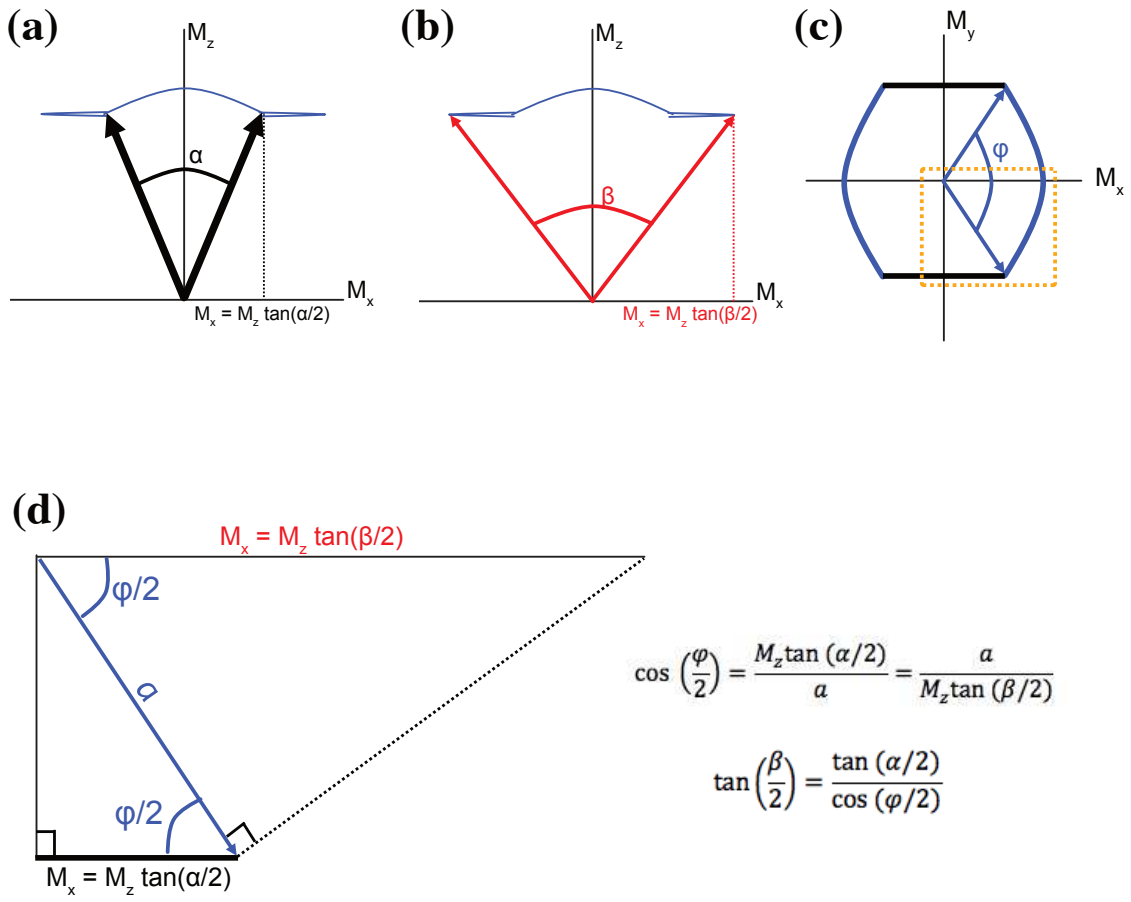


Figure 2.6: The magnitude of transverse magnetization along the x-axis, $M_x$, for the flip angle $\alpha$ in (a), and the "effective flip angle", $\beta$, in (b), which takes into account the phase-offset caused by precession of spins oscillating off-resonance. The angle of precession, $\phi$, is shown on the $M_x$-$M_y$ axis in (c). The dotted yellow box in (c) is expanded on in panel (d), which shows the relationship between $\alpha$, $\beta$, and $\phi$.

The signal defined by Equation 2.23, with and without precession effects, is demonstrated in Figure 2.7.

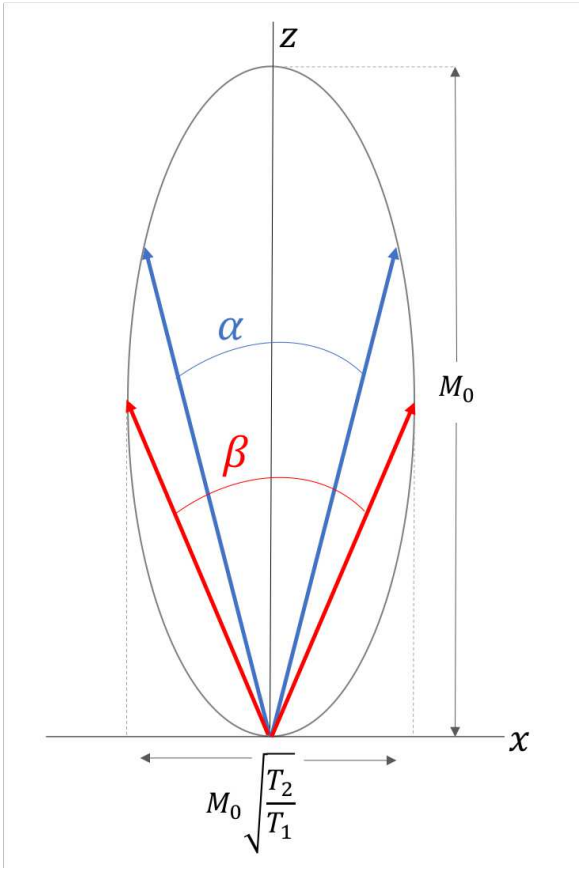


Figure 2.7: The transverse component of steady-state magnetization defines an ellipse with height and width demonstrated here. The choice of flip angle, $\alpha$, defines where along the ellipse the signal is measured. Off-resonance effects will alter this position, resulting in an "effective flip angle" of $\beta$.[58]

We can now consider the transverse component of magnetization generated by using a flip angle of $\beta$. Letting $M_{xy}^2 = M_x^2 + M_y^2$, and substituting $M_{xy} = M\sin(\beta/2)$, and $M_z = M\cos(\beta/2)$ in Equation 2.23, we can achieve

$$M_{xy} = \frac{M_0}{\cot(\beta/2) + \frac{T_1}{T_2}\tan(\beta/2)} \tag{2.25}$$

A full derivation of Equations 2.23 and 2.25 is provided in Appendix A.

As expected, Equation 2.25 demonstrates that the contrast in bSSFP acquisitions is not a simple $T_1$ or $T_2$ weighting. Instead, it depends on the difference of $T_1/T_2$ ratios between tissues. Importantly, the signal strength shows a strong dependence on the effective flip angle, $\beta$; for given $T_1$ and $T_2$, the optimal flip angle, $\beta_{opt}$, would

bring $M$ to touch the ellipse at its maximum width, such that the peak signal is given by $M_{peak} = \frac{1}{2}M_0\sqrt{\frac{T_1}{T_2}}$. For similar $T_1$ and $T_2$, the maximum attainable signal therefore approaches $\frac{1}{2}M_0$. It is this feature that makes bSSFP sequences the most SNR efficient - it is possible to continuously acquire a signal of $\frac{1}{2}M_0$.[42]

For a fixed flip angle, $\alpha$, an increase in off-resonance precession, $\phi$, towards 180° leads to a corresponding increase in $\beta/2$ towards 90° (see Equation 2.24), causing the signal defined by Equation 2.25 to approach zero. This relationship between precession angle and measured signal is demonstrated in Figure 2.8.

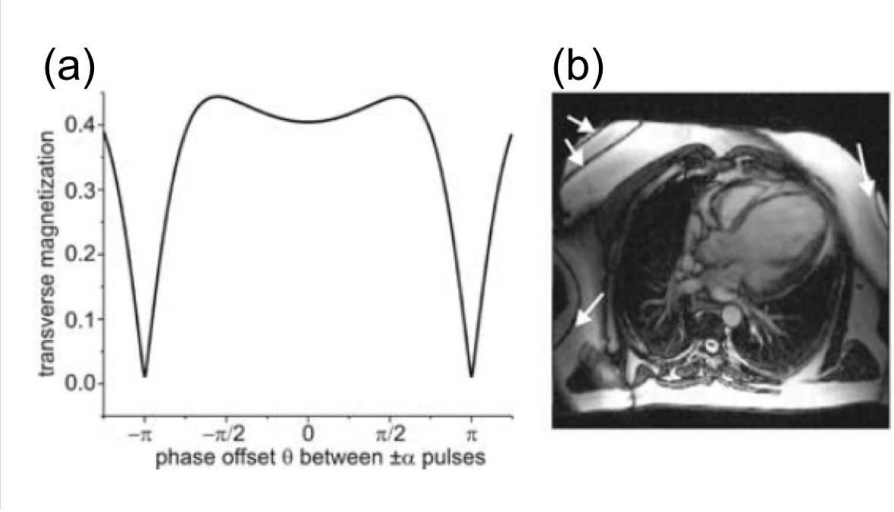


Figure 2.8: (a) Steady-state transverse magnetization achieved with varying phase offsets, denoted here as $\theta$. Notably, for this example, the signal increases initially as the offset deviates from zero. This is due to the flip angle used being less than $\beta_{opt}$ such that the signal lays above the maximum width of the ellipse depicted in Figure 2.7 - offsets increase the angle between $M$ and the $z$-axis, thereby bringing the magnetization vector down the edges of the ellipse to its maximum width. The signal dropout that occurs with $\theta = 180°$ is shown by the white arrows in panel (b). The figure is a modified version of that displayed by Scheffler and Lehnhardt 2003.[42]

The phase offset, $\phi$, accrued by a spin over time TR is simply its angular frequency, $2 \cdot 180° \cdot \nu$, multiplied by TR such that $\phi = 2 \cdot 180° \cdot \nu \cdot \mathrm{TR}$. Therefore, spins that are 180° out of phase after TR must be oscillating with a frequency of

$$\begin{aligned} \theta &= 2 \cdot 180° \cdot \nu \cdot \mathrm{TR} \\ \pm 180° &= 2 \cdot 180° \cdot \nu \cdot \mathrm{TR} \\ \nu &= \frac{1}{\pm 2TR} \end{aligned} \tag{2.26}$$

Thus, dark bands occur in image locations where spins are off-resonance by frequencies of $\nu = \pm\frac{1}{2TR}$ (Panel (b) of Figure 2.8), and a precession frequency "pass-band" of width $1/TR$ exists. As discussed, off-resonance can be a result of inhomogeneities in $B_0$, or of susceptibility differences between tissues. Fortunately, Equation 2.26 shows that a decrease in TR will reduce the amount of time for dephasing, requiring very high frequency offsets for bSSFP banding artifacts to be produced. This feature of bSSFP acquisitions can be considered when imaging regions that typically exhibit high likelihood of susceptibility-induced distortions.

### Phase Cycling

When a reduction in TR is not sufficient, or not possible due to SAR constraints, *phase cycling* of the RF pulse can be used to minimize the presence of off-resonance banding artifacts. This method requires the collection of two or more separate data sets, and for the sake of simplicity, the case of two acquired data sets is considered here; in this case, one acquisition uses RF pulses that have alternating phases (as depicted in Figure 2.5), and one in which the phase of the RF pulse is not altered.

In the first case, spins with $\phi = 180°$ after TR do not contribute to the steady-state signal, as described previously. In the second case, spins with $\phi = 180°$ do contribute to the steady-state signal, and instead, those that are on resonance (i.e. $\phi = 0°$) do not. This has the effect of shifting the locations of banding artifacts in the image to regions where spins are on-resonance, as demonstrated in Figure 2.9.

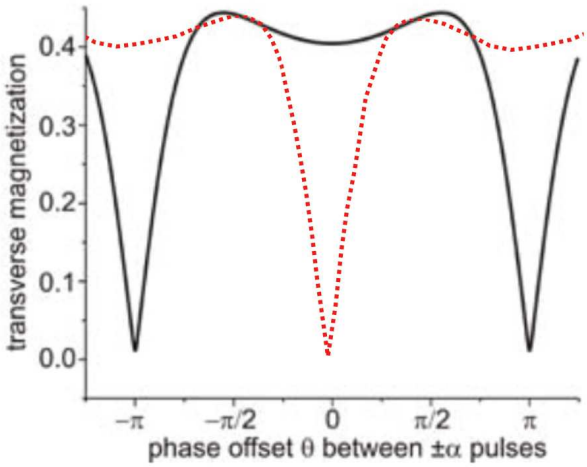


Figure 2.9: Cycling the phase of the RF pulse used in a bSSFP acquisition shifts the banding artifacts to locations where spins are oscillating on-resonance, as indicated by the red dotted line. The image is a modified version of that presented by Scheffler and Lehnhardt 2003.[42]

The two images can then be combined so as to minimize the presence of artifact. Numerous methods have been proposed to combine the two data sets, and the most commonly used is that of maximum intensity projection (MIP).[60] MIP reconstruction calculates a new 3D image by using the maximum of each data set on a pixel-wise basis. Brown et al.[61] demonstrate that the SNR of an image resulting from MIP increases with the number of images included in the reconstruction. The trade-off, however, is increased scan time which can lead to blurring artifacts due to patient motion between acquisitions. In addition, the signal variation, as seen in Figure 2.9, can never be fully removed from all tissues.

## 2.3 Super Resolution

Apart from the decrease in image quality, low resolution (LR) acquisitions confer a number of advantages: decreased acquisition times, reduced patient heating, and reduced patient motion. This approach applied to bSSFP acquisitions has the additional advantage of increased artifact tolerance due to a decrease in TR. The application of a neural network, trained to perform the task of super resolution, could allow for the advantages of LR bSSFP imaging to be realized, without significantly degrading image quality.

### 2.3.1 Neural Networks

In general, a neural network consists of numerous connected computational units known as neurons. When working with image data, each pixel in the image corresponds to a neuron in the network's input layer. Subsequent neurons are organized into layers such that input data enters the input layer, makes its way through the computational elements of the hidden layers, and exits the model at the output layer. Figure 2.10 depicts a simple neural network consisting of two hidden layers. The number and type of hidden layers used in a model, and how they are connected, defines a model's *architecture*.

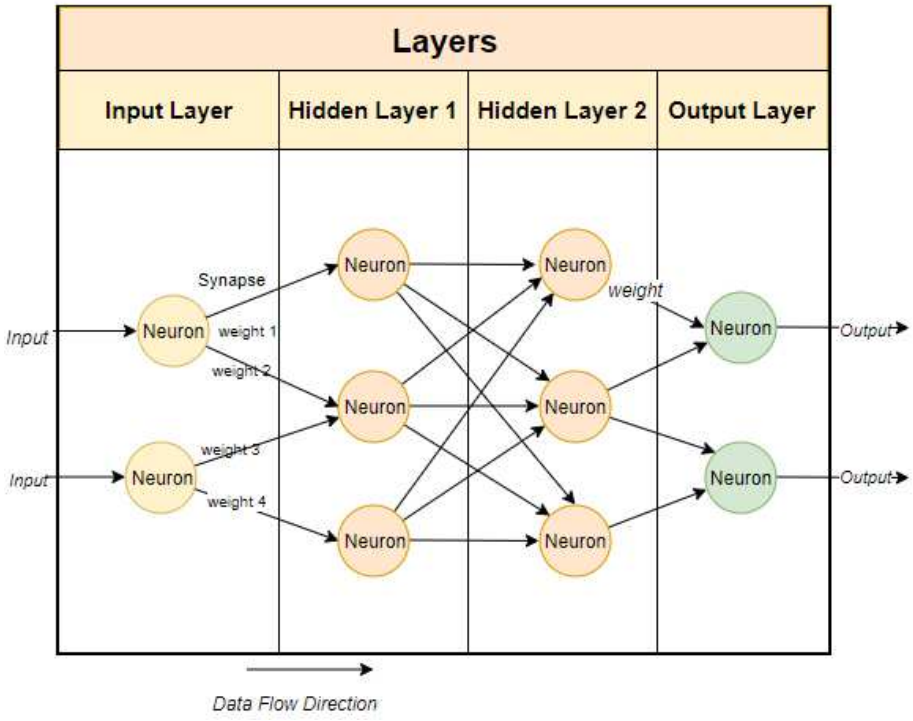


Figure 2.10: The basic components of a neural network organized into input, hidden, and output layers. The Figure is reproduced from Malik 2019.[62]

The *weight* of a neuron's contribution to the neurons in the subsequent layer is an important feature of a neural network. They are traditionally initialized as small, random numbers, and it is these values that are fine-tuned during the training process such that the desired task can be performed.

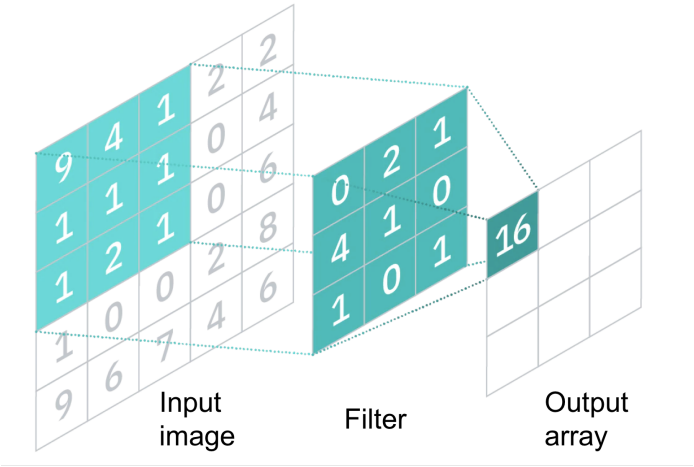


Figure 2.11: A single convolutional filter (i.e. $W \times H \times C = 3 \times 3 \times 1$) acts on an input image of size $5 \times 5 \times 1$, using a stride of 1, producing an output array of size $3 \times 3 \times 1$. Convolutional layers in a CNN can consist of multiple filters, producing multiple output layers that capture various high-level features in the input image. The figure is a modified version of that produced by IBM Cloud Education.[64]

### 2.3.2 Fundamental Layers of Convolutional Neural Networks

Convolutional neural networks (CNNs) are capable of extracting high-level features from an image and are, therefore, very commonly used for image-oriented tasks.[63]

The fundamental building block of a CNN is the convolutional layer, which utilizes a *convolutional filter*. The convolutional filter is commonly a tensor of shape $H_j \times W_j \times C_j$ where $H_j$ and $W_j$ are the height and width of the filter at layer $j$, and $C_j$ is the number of channels, or feature maps, to be extracted at the $j$th layer. As the convolutional filter is generally much smaller than the input it is being applied to, multiple steps are required when performing the convolution operation. The size of the step taken is termed the *stride* and is typically set to 1 or 2. The example in Figure 2.11 demonstrates a convolution with a stride of 1, which results in the output array having smaller dimensions than the input. If this is undesirable, zero padding along the borders of the input can be performed such that the convolution operation does not change the dimensions of the input. Importantly, the values that make up each filter are the weights that need to be trained in a CNN. The size and number of filters used at a given convolution layer are an example of a model's hyperparameters and can be chosen to reflect the number and size of features in the image that are to be extracted.

Following a convolutional layer is an *activation layer*, which allows the network to learn non-linear relationships between the input and output. The most common activation function is the Rectified Linear Unit, ReLU, defined as $A_{ReLU}(z) = \max(0, z)$ for input $z$. The LeakyReLU function is a modified version of ReLU, designed to prevent neurons from being assigned a value of zero over numerous updates to the weights. Leaky ReLU returns $\alpha z$ (where $\alpha$ is typically chosen to be 0.01) when $z < 0$, and $z$ when $z > 0$.[65] The sigmoid activation function is commonly used when the network is trained to perform a binary classification task. The function, $A_{sigmoid}(z) = 1/(1 + e^{-z})$, returns values between 0 and 1, and is therefore ideal for determining the probability of the input belonging to a given class.[63] Activation functions are applied in an element-wise manner and therefore do not alter the dimensions of the input. Notably, activation functions do not contain any trainable parameters.

*Batch normalization*[66] is a procedure that can be performed before or after a convolutional layer. Here, "batch" refers to a subset of the training data that is provided to the network before the weights are updated. To prevent the distribution of weights at a given layer from changing between each update, the output of the previous layer is normalized to a standard distribution by subtracting the batch mean and dividing by its standard deviation. Batch normalization has been said to speed up training and reduce the dependency on weight initialization.[63]

Fully connected, or *dense* layers are typically used when a classification task is being performed. A dense layer is one in which the layer's neurons are each connected to every neuron in the preceding layer, with each connection having an associated weight used to determine its contribution. In the case of binary classification, the final layer consists of a single, fully connected neuron, whose value can be input to a sigmoid activation function, for example, to determine the final class to which the input is assigned.

### 2.3.3 Model Architectures

Many networks have been developed that make use of CNNs, and an overview of some of the common architectures is provided by Lundervold et al.[63] Of relevance to the work presented in this thesis is the Residual Network (ResNet) and the Generative Adversarial Network (GAN).

A Residual Network (ResNet), is a CNN that employs many "skip connections", which do not have trainable parameters, but simply add the output of a previous layer to a layer ahead. The *residual blocks* created are the fundamental component of a ResNet[22] (Figure 2.12(a)). The connections between layers will include a linear projection, if necessary, to ensure the layers' outputs have identical dimensions before being combined. Figure 2.12 also demonstrates a *dense residual block*, where the connections are such that the output of one layer is concatenated with the output from numerous subsequent layers. This connection between upper and lower layers of a network allows the preservation of information from layers higher up in the architecture. Compared to CNNs, ResNets have the advantage of preserving higher level information without the need to train additional parameters.

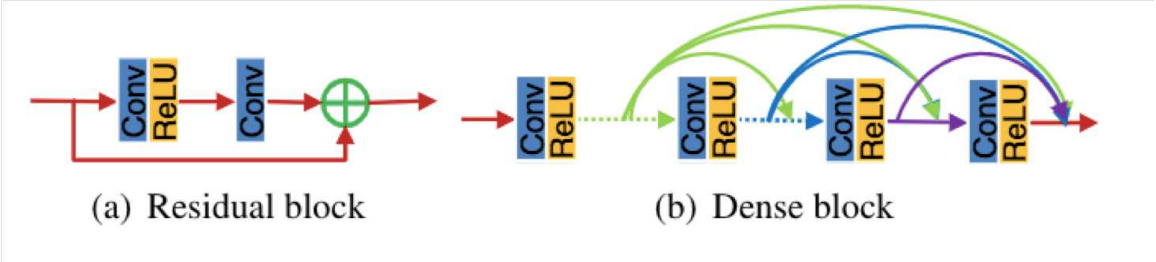


Figure 2.12: (a) A diagram of a residual block where the output of a shallow layer (not shown) is added to the output of a deeper convolutional layer. A residual dense block (b), in which each the output of a convolutional layer (followed by ReLU activation) is concatenated to the output of each subsequent activation. The figure is reproduced from Zhang et al. 2018.[67]

A generative adversarial network (GAN)[68] consists of two sub models - a *generator* and a *discriminator*. When performing image-related tasks, the generator learns to produce a modified version of the input image, and the discriminator attempts to classify the image as originating from the domain of "real" images, provided during training, or as "fake", i.e. produced by the generator. The overall architecture of a GAN is shown in Figure 2.13. The generator model typically consists of a CNN such as ResNet, as discussed. The discriminator can also be comprised of a CNN, but must include fully-connected layers, and an activation layer such as the sigmoid function, as the last computational steps. This structure permits the discriminator to output one final variable: the probability of an image belonging to the real image domain. Training a GAN requires optimizing a min-max problem - the probability of

the discriminator correctly labelling an image as real or fake is maximized while the generator is trained so as to minimize the probability of the discriminator labelling its images as fake. This encourages the generator to produce images that look more and more like those in the real image domain. The training therefore takes place in an alternating manner until the discriminator is unable to determine what domain an image originates from (i.e. it predicts equal probabilities for fake and real images). Further discussion on training of GANs is provided in the following section.

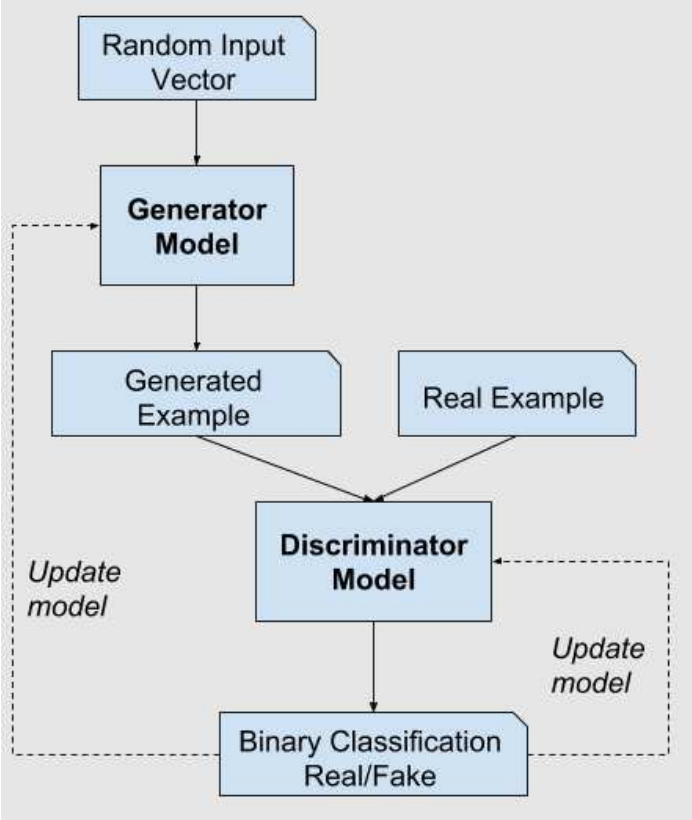


Figure 2.13: An overview of the architecture of a generative adversarial network (GAN). The figure is reproduced from Brownlee 2017.[69]

### 2.3.4 Training, Validation, and Testing

Models can be trained via *supervised* or *unsupervised learning*. In the context of performing the super resolution task, supervised learning is feasible when paired examples of low resolution (LR) input and high resolution (HR) target images are available to train the model on. During training, the model is sequentially shown a large number of LR-HR training pairs (or multiple pairs in the form of a batch). It then gradually

learns the relationship between an LR input image and its corresponding HR target via optimization of an objective function which compares the model's output to the target. When paired examples are not available, a model can undergo unsupervised learning. In this case, the model tries to generate an image, based on LR input, that fits into the domain of HR images that it is provided. Further discussion of possible optimization functions is provided in in Section 2.3.5. The iterative optimization process considers the effect of each of the network's weights as it adjusts them and attempts to find the optimal solution to the multidimensional objective function. The mathematical process of doing so is known as *backpropagation*.

In parallel to the training process, the model can be fed previously unseen images to generate a prediction based on its weights at that point in training. Doing so is a way to estimate of the model's current skill and is known as *validation*. The skill level can be quantified via calculation of the loss function or via a new function more relevant to the desired outcome. When learning the SR task via supervised learning, for example, validation be completed via calculation of the peak signal-to-noise ratio (Section 2.3.5) between the model's prediction and the HR target. Validation is not used to adjust model weights, but can instead be used to tune *hyper-parameters* of the model such as the number and type of layers.

Once training is completed, the final model can be used to make predictions based on unseen test data in a stage known as testing.

Training, validation, and testing of a neural network can require very large datasets. Ledig et al.,[24] for example, trained their network, SRGAN, on 350 thousand images from the ImageNet[70] database of natural images. In the context of MRI, such large data sets are difficult to acquire. For this reason, multiple publicly available MR data sets have been developed to encourage model development and training (see Bento et al.[71] for a summary of some of the publicly available brain imaging data sets). It is important to note, however, that such data sets, while valuable, may not reflect the characteristics of images acquired on a particular MR system, and may lead to misleading or unreliable results.[71]

### 2.3.5 Objective Functions

The choice of objective function is crucial for both supervised and unsupervised learning methods, as optimization of the function is what determines the adjustment of model weights.

For image tasks, supervised learning techniques most commonly attempt to minimize the pixel-wise mean squared error (MSE) as it is a convex and differentiable function. For an image, $Y$, output by a model, and the reference image, $X$, both with $N$ pixels, the mean squared error is given by:

$$MSE(X,Y) = \frac{\sum_{i=1}^{N}(X_i - Y_i)^2}{N} \tag{2.27}$$

The peak signal-to-noise ratio, PSNR, is related to MSE, and is also a commonly used metric for image-related tasks. If we consider the output image to be equivalent to the reference image with additional undesirable features, i.e. "noise", the difference between predicted and reference images results in the "noise" signal. MSE is therefore the power of the average noise signal. As such, PSNR is a measure of the ratio of the maximum signal in the reference image to the average noise signal:

$$PSNR(X,Y) = 10 \cdot log_{10}\left(\frac{MAX_X^2}{MSE(X,Y)}\right) \tag{2.28}$$

However, it is widely accepted that MSE, and consequently PSNR, do not correlate well with humans' perception of image quality. MSE, for example, assumes that the effect of noise is independent of local image characteristics, whereas the human visual system's sensitivity to noise depends on local luminance, contrast, and structure.[72]

Introduced in 2016, the perceptual loss function calculates the average MSE between higher level features of the images, extracted from a CNN known as VGG (named after the Visual Geometry Group that developed it).[73] This type of loss function better reflects human perception and was shown to produce predictions that out-perform identical models trained on MSE.[74] For the $j$th layer of the VGG network, a feature map of shape $C_j \times H_j \times W_j$ is created, and the perceptual loss is then given by

$$L_{perceptual}^{\mathrm{VGG},j}(X,Y) = \frac{1}{C_j H_j W_j} ||\mathrm{VGG}_j(X) - \mathrm{VGG}_j(Y)||^2 \tag{2.29}$$

where $\mathrm{VGG}_j$ is the output of the $j$th layer of the VGG network.

Training a GAN requires optimization of the max-min adversarial problem defined by Goodfellow et al.;[68] for a generator, $G$, and discriminator, $D$ acting on fake images, $z$, and real images, $x$, this problem can be defined as

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}(x)}[logD(x)] + \mathbb{E}_{z\sim p_z(z)}[log(1-D(G(z)))] \tag{2.30}$$

where $p_{data}(x)$, and $p_z(z)$ represent an image, $x$, drawn from the distribution of real images $p_{data}$, and a fake image, $z$, drawn from the distribution of images created by the generator, $p_z$, respectively.

## 2.4 Experimental Details

To accomplish the research objectives outlined in Section 1.6, images resulting from two unique 0.5T bSSFP protocols, one designed to examine the temporal bone and one to image the paranasal sinuses, were acquired. What follows is a description of the MR system used and the protocols developed.

### 2.4.1 Hardware

Imaging protocols were developed on a 0.5T head-only MR system (Synaptive Medical, Toronto, Canada) designed for point-of-care neuroimaging. The system consists of a cryogen-free, closed conduction cooled superconducting magnet, and high performance asymmetric gradient coils with a peak strength of 100 mT/m and peak slew rate of 400 T/m/s. An adjustable 16 channel head-coil enables patient-specific placement, and is combined with a digital receive chain.

### 2.4.2 Temporal Bone Protocol

The complex region of the inner ear requires high resolution to visualize effectively, and as such, an acquisition strategy with isotropic resolution of 0.3 mm was prioritized

when developing the temporal bone protocol. A maximum gradient strength of 60 mT/m permitted a TR of 6.7 ms.

With these parameters, however, it was found that some subjects' images demonstrated band-like artifacts that traversed the FOV and often obscured the relevant anatomical structures. These artifacts were confirmed to be due to concomitant gradient fields by altering the placement of the participant on the scanning bed. The artifacts remained inferior to the inner ear structures unless the participant was situated inferior to the isocenter of the gradient fields. Since ideal placement of participants in the scanner was not always possible, the protocol was modified to use a maximum gradient strength of 30 mT/m so as to reduce the effect of concomitant fields and to allow for more flexibility of participant positioning. This led to an increase of TR to 7.0 ms, and an increase in scan time of 19 s. All other imaging parameters are listed in Table 2.1.

| Imaging Parameter | |
|---|---|
| Field of view (cm) | 18 |
| Acquired isotropic resolution (mm) | 0.6 |
| Number of slices | 164 |
| Number of excitations | 1 |
| Receive Bandwidth (kHz) | 70 |
| Repetition time (ms) | 6.7 / 7.0 |
| Echo time (ms) | 3.3 |
| Flip angle (°) | 60 |
| Interpolated isotropic resolution (mm) | 0.3 |
| Maximum gradient strength (mT/m) | 60 / 30 |
| Slew rate (T/m/s) | 200 |
| Scan time | 4min. - 4min. 19s. |

Table 2.1: Imaging parameters for the 0.5T bSSFP temporal bone protocol.

An image with concomitant gradient artifacts, acquired with a maximum gradient strength of 60 mT/m, is shown in Figure 2.14, along with an image of the same patient acquired with a maximum gradient strength of 30 mT/m.

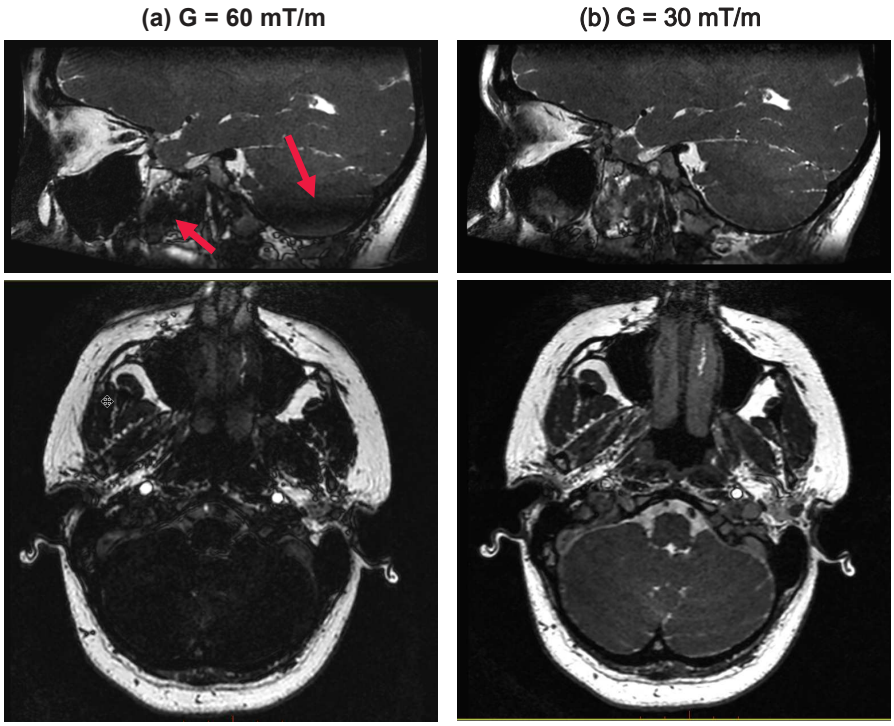


Figure 2.14: Images acquired from a healthy volunteer with the 0.5T bSSFP temporal bone protocol using a maximum gradient strength of (a) 60 mT/m, and (b) 30 mT/m. The red arrows in the top panel of (a) indicate the loss of signal resulting from concomitant gradient fields, and the bottom panels show the image acquired at this axial location. The volunteer's position along the superior-inferior direction was unchanged between acquisitions.

The temporal bone images in this work were acquired from a total of 38 participants. The initial version of the low-field protocol was used to acquire images from the first 10 participants, with the remaining 28 being imaged with the modified protocol. Recruitment and acquisition protocols were approved by the Nova Scotia Health (NSH) Research Ethics Board (REB), and participants' oral and written consent were obtained. The protocol also provided permission to access any participants' retrospectively acquired clinical scans, making research objectives 1 and 2 feasible (Section 1.6).

### 2.4.3 Sinus Protocol

The multiple air/tissue and air/bone interfaces in the paranasal sinuses exhibit inherently high contrast but have large susceptibility differences that cause off-resonance

effects. The resulting banding artifacts can mimic drainage pathways or obscure the fine structures that separate the air-filled regions. For this reason, a small TR value of 3.5ms, enabled by high peak gradient strength and slew rate, was prioritized in our sinus protocol in order to reduce off-resonance effects, and ultimately minimize the presence of banding artifacts in the anatomical structures of interest. As such, a small flip angle of 30° was used to maintain contrast. The imaging parameters for the 3D bSSFP sinus protocol are shown in Table 2.2.

| Imaging Parameter | |
|---|---|
| Field of view (cm) | 18 |
| Acquired isotropic resolution (mm) | 0.8 |
| Number of slices | 176 |
| Number of excitations | 1 |
| Receive Bandwidth (kHz) | 150 |
| Repetition time (ms) | 3.5 |
| Echo time (ms) | 1.6 |
| Flip angle (°) | 30 |
| Interpolated isotropic resolution (mm) | 0.4 |
| Maximum gradient strength (mT/m) | 50 |
| Slew rate (T/m/s) | 200 |
| Scan time | 1min. 52s. |

Table 2.2: Imaging parameters for the 0.5T bSSFP paranasal sinus protocol.

# Chapter 3

# Artifact-resistant balanced steady-state free precession imaging of the temporal bone and paranasal sinuses without phase-cycling at 0.5T

## 3.1 Study Objectives and Hypotheses

The primary objective of this study is to perform a quantitative analysis of the artifact tolerance afforded by the 0.5T temporal bone and paranasal sinus protocols by tabulating the number of images with artifacts traversing relevant structures. A theoretical decrease in susceptibility-induced off-resonance effects at 0.5T, compared to conventional field strengths, leads us to hypothesize that:

1. the majority (>50%) of images from the 0.5T temporal bone and paranasal sinus protocols will be free of artifact traversing relevant structures.

Since our institution's clinical temporal bone exams include bSSFP images, we will compare the artifact tolerance between 0.5T and these clinical exams. The clinical temporal bone images are obtained at 3T with phase-cycled bSSFP acquisitions reconstructed with MIP. As such, we hypothesize that:

2. the number of temporal bone images without banding artifacts traversing relevant structures in images acquired at 0.5T will be less than those acquired clinically.

Provided sufficient artifact tolerance, the contrast-to-noise (CNR) ratio between neighbouring tissues is a key metric in satisfactory visualization.[75] The majority of the structures examined herein abut air-filled spaces, and the SNR of each structure is therefore equivalent to a measure of CNR. A study of radiologist preference, performed by Owen et al.[76] yielded the result that an SNR of $\approx 20$ is desirable, with any excess being traded for resolution. To that end, the SNR of relevant structures will be

calculated for each protocol. Given the decrease in $T_1$ values at low field, and the use of fast gradients and a conformable head-coil, we hypothesize that:

3. the SNR in images resulting from the 0.5T temporal bone and paranasal sinus protocols will be the same as, or greater than that preferred for visualization by radiologists (i.e. SNR $\geq$ 20) for all structures examined.

## 3.2 Methods

### 3.2.1 Image Acquisition

#### Temporal Bone

A subset of temporal bone images from 21 participants were included in this study. The inclusion criteria were that participants' clinical exams were performed at 3T, and that their 0.5T scans were acquired under the newest version of the temporal bone protocol (i.e. TR = 7.0ms and maximum gradient strength 30 mT/m, with all other imaging parameters as listed in Table 2.1). The cohort included 9 participants that identified as male (mean age 52 years $\pm$ 11) and 12 participants that identified as female (mean age 48 years $\pm$ 13). Combined cohort mean age was 52 years $\pm$ 12 years ranging from 22 to 66 years.

Participants' SOC temporal bone images were acquired at 3T on the Discovery MR750 system with a rigid, 32-channel head coil (General Electric Healthcare). All SOC images were produced from phase-cycled bSSFP acquisitions. Phase cycling consisted of one acquisition with a consistent RF pulse phase, and a second acquisition with an RF pulse phase that alternated between 0° and 180°. The final image was reconstructed with MIP. This cohort of images will henceforth be referred to by its vendor-specific name of phase-cycle Fast Imaging Employing Steady-state Acquisition, i.e. FIESTA-C. The acquisition parameters for the 3T FIESTA-C images were chosen to be the best for SOC imaging of the temporal bone at 3T by our institution, and are listed in Table 3.1.

To quantify the effect of field strength on artifact tolerance, three additional bSSFP images were acquired from a healthy volunteer (48-year-old male) under an REB approved protocol for technique development. Two of the images were acquired on the 3T system - one image was acquired with the same parameters as the clinical

| Imaging Parameter | |
|---|---|
| Field of view (cm) | 18 |
| Acquired isotropic resolution (mm) | 0.6 |
| Number of slices | 136 or 144 |
| Number of excitations | 2 |
| Receive Bandwidth (kHz) | 73 |
| Repetition time (ms) | 5.5 - 5.6 |
| Echo time (ms) | 2.2 - 2.3 |
| Flip angle (°) | 60 |
| Number of phase cycles | 2 |
| Interpolated isotropic resolution (mm) | 0.3 |
| Maximum gradient strength (mT/m) | 55 |
| Slew rate (T/m/s) | 140 |
| Scan time | 5min. 22s - 5min. 45s. |

Table 3.1: Imaging parameters used for clinical 3T 3D axial FIESTA-C acquisitions.

FIESTA-C protocol (NEX = 2, phase-cycling and MIP), and a second was acquired without phase cycling methods (NEX = 2), and therefore MIP was not performed (scan time = 4 min 6 s, and all other parameters as listed in Table 3.1). The third image was collected on the 0.5T system with the parameters shown in Table 2.1 (i.e. NEX = 1, no phase cycling).

**Paranasal Sinus**

Under our REB approved paranasal sinus protocol, 25 participants were scanned who had clinical requisitions for suspected sinus abnormalities requiring CT imaging. Informed written consent was obtained. The sinus cohort included 11 participants that identified as male (mean age 55 years ± 13) and 14 participants that identified as female (mean age 57 years ± 12). The combined cohort mean age was 56 years ± 12, ranging from 33 to 73 years.

Although clinical MRI of the sinuses is not performed at our institution, a 3T protocol was developed under the REB-approved technique development protocol, with parameters similar to that of the 0.5T protocol (FOV=18cm, matrix = 226×226, 172 slices, isotropic acquired resolution = 0.8mm, NEX=1, RBW = 73 kHz, TR/TE = 4.2ms / 1.6ms, flip angle = 30°, interpolated isotropic resolution = 0.4mm, scan time = 2 min 48s, no phase cycling) to exemplify the challenges of imaging the paranasal

sinuses at conventional field strengths. One healthy volunteer (48-year-old male) was scanned with this protocol, as well as the 0.5T paranasal sinus protocol (parameters shown in Table 2.2) to facilitate comparison.

### 3.2.2 Image Analysis

To quantify the artifact tolerance of the 0.5T temporal bone protocol (NEX = 1, no phase cycling or MIP) in comparison to the clinical 3T FIESTA-C acquisitions (NEX = 2, phase cycling and MIP), as well as the artifact tolerance of the 0.5T paranasal sinus protocol (NEX = 1, no phase cycling), a researcher was first trained by one radiologist on how to identify problematic artifacts. Next, all images (excluding those of the healthy volunteer) were examined individually and the presence of any banding artifacts in the structures of interest were recorded and tabulated.

To calculate the SNR values resulting from the low-field protocols, two board-certified neuroradiologists provided guidance on the structures critical to characterize under each protocol, and regions of interest (ROIs) were drawn around them. For the temporal bone exam, this included the left and right vestibules, cochlea, and superior semi-circular canals, and regions of cerebral spinal fluid (CSF) surrounding the left and right facial, cochlear, and vestibular nerves. Additionally, a white matter region of the brain, and the fourth ventricle were included in the analysis. For the sinus images, ROIs were drawn around a region of CSF superior to the cribriform plate, inside one globe, inside one maxillary sinus, and mucosa in one of the middle nasal turbinates. All ROIs were drawn in the plane which best visualized the structure and spanned a minimum of 5 slices. A noise region was drawn in each image outside the head along the frequency-encode direction.

The average SNR for each anatomical structure was calculated according to

$$SNR = 0.70\frac{S}{SD_{air}} \tag{3.1}$$

where $S$ is the mean signal in the ROI and $SD_{air}$ is the standard deviation of signal in the noise ROI. It has been shown that the transformation from complex MR data to magnitude image results in the distribution of pixel intensities, in the presence of Gaussian noise, following the Rician probability distribution function (PDF). For regions in which there is high signal intensity, the PDF behaves like a Gaussian PDF,

while reducing to a Rayleigh PDF in the absence of signal. In addition, the use of a 16 channel RF receive coil alters the Rayleigh noise distribution. Combined, these effects lead to a correction factor of 0.70 being applied to compare ROIs containing signal to those of background noise.[77]

Once calculated, the SNR for left and right temporal bone structures were concatenated such that there was one set of results for each structure at each field strength. All SNR calculations were done for all subjects and the mean and 95% confidence interval across subjects were determined for each structure.

It is important to note that the degree to which the MIP method alters the distribution of pixel intensities differs from one PDF to another. We have empirically confirmed that performing a MIP operation on multiple phase-cycled images significantly alters signal and noise distributions of the resulting image, in a way that is not as easily corrected as standard Fourier reconstructions. Despite quantitative comparisons of MIP and other acquisition techniques being performed in other studies, there is no consensus in the literature on how to calculate SNR in MIP reconstructions. Thus, we believe that a direct comparison of SNR between the 0.5T bSSFP (without phase cycling and MIP) and 3T FIESTA-C acquisitions (with phase cycling and MIP) may not be meaningful.

## 3.3 Results

### 3.3.1 Temporal Bone

Representative images from temporal bone acquisitions at 0.5T (NEX = 1, no phase cycling) and 3T (NEX = 2, phase cycling and MIP) are shown in Figure 3.1. The additional 3T acquisitions are shown in Figure 3.2, along with the 0.5T acquisition of the same healthy volunteer. Together, these images portray the necessity of phase cycling and MIP when acquiring temporal bone images at 3T.

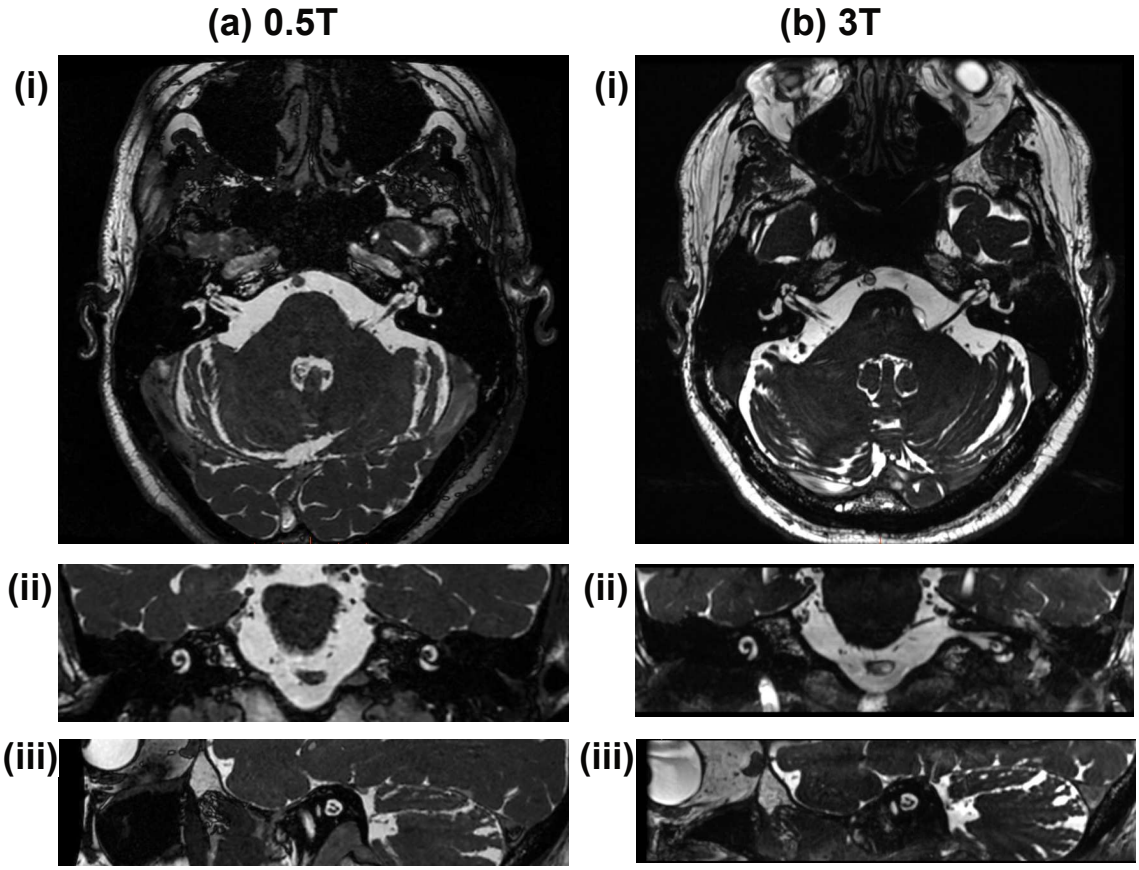


Figure 3.1: Axial (i), coronal (ii), and sagittal (iii) views of representative (a) 0.5T bSSFP and (b) 3T FIESTA-C temporal bone acquisitions. 3T data are displayed as a MIP of phase-cycled acquisitions. 0.5T data were acquired with NEX = 1, while 3T data were acquired with NEX =2 and displayed as a MIP of the phase-cycled acquisitions. The 3T images were acquired with a repetition time of 5.5ms and an echo time of 2.2ms, while all other parameters are as those listed in Table 3.1.

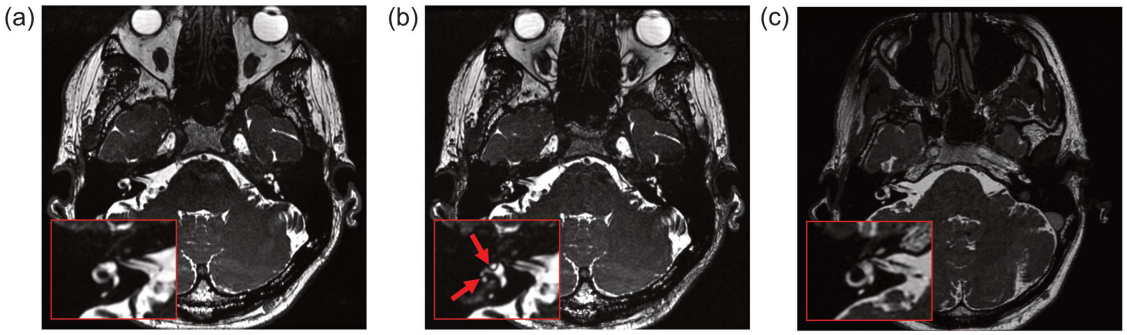


Figure 3.2: Images resulting from (a) FIESTA-C at 3T (NEX = 2, phase cycling and MIP), (b) FIESTA at 3T (NEX = 2, no phase cycling and no MIP), and (c) bSSFP (NEX = 1, no phase cycling) at 0.5T. The red boxes are zoomed-in views of the region surrounding the right vestibule and lateral SSC. Red arrows in (b) point to banding artifacts that are not present at similar anatomical locations in (a) and (c).

Of the 21 sets of temporal bone images examined, 4 (19%) demonstrated banding artifacts in the structures of interest at 0.5T (NEX = 1, no phase cycling), while 18 (86%) of the 3T FIESTA-C (NEX = 2, phase cycling and MIP) acquisitions showed banding. An example image with artifacts present at 3T is shown in Figure 3.3 along with the 0.5T image of the same participant.

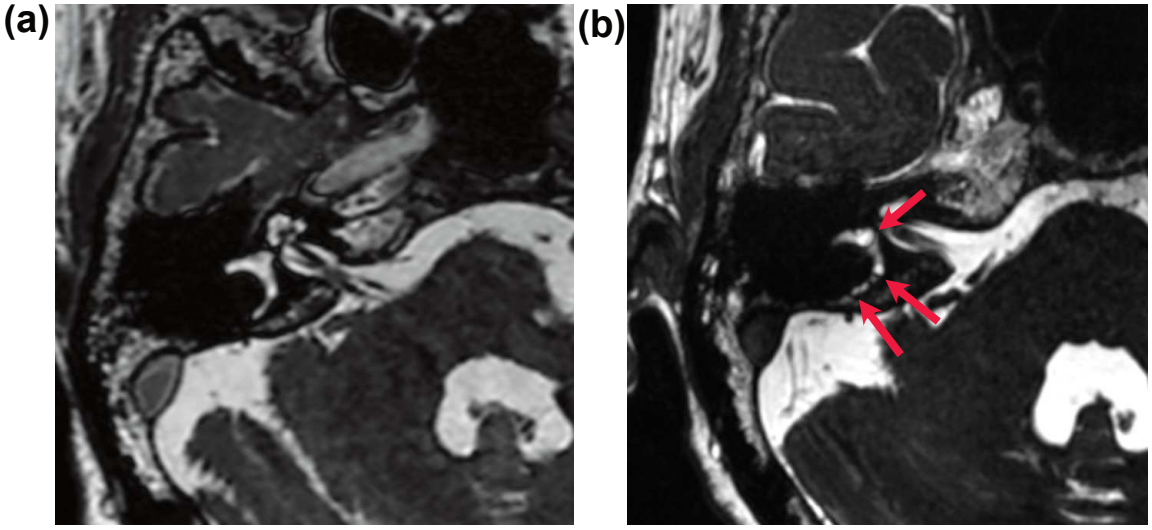


Figure 3.3: Images resulting from a (a) 0.5T bSSFP (NEX = 1, no phase cycling) acquisition and (b) a 3T FIESTA-C (NEX = 2, phase cycling and MIP) acquisition. The red arrows in (b) indicate banding artifacts in the right vestibule and lateral semicircular canal, which are not present in the 0.5T acquisition.

The ROIs examined in the 0.5T temporal bone acquisitions are depicted in Figure 3.4 and the resulting SNR achieved are displayed in Table 3.2.

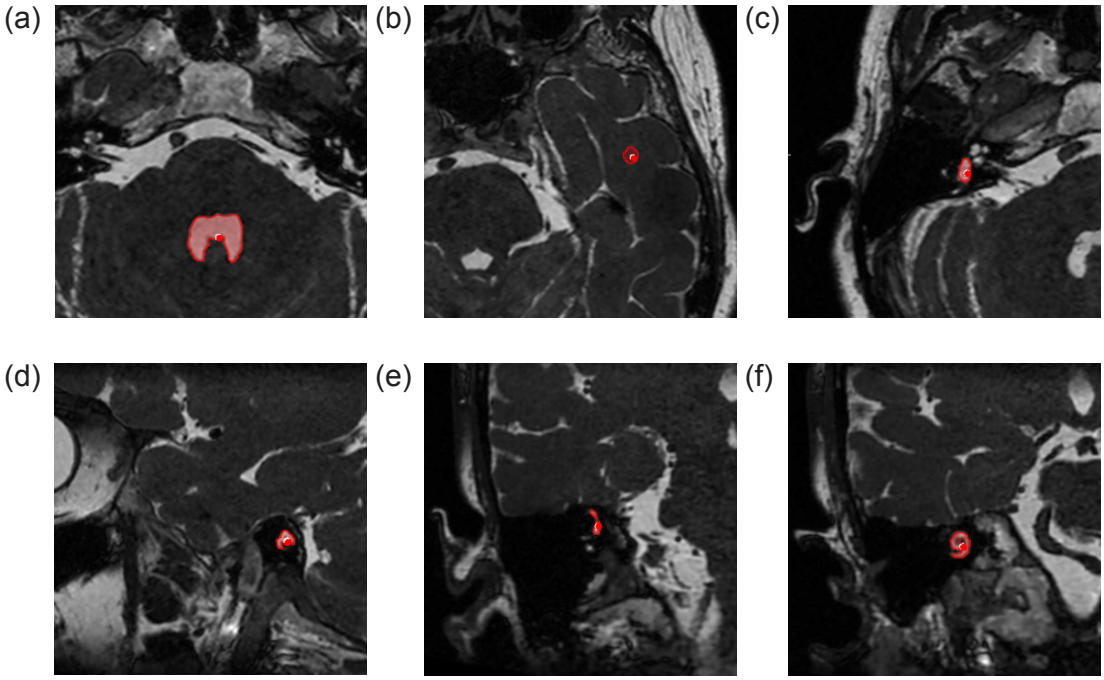


Figure 3.4: The ROIs examined under the 0.5T (NEX = 1, no phase cycling) temporal bone protocol are shown in red: (a), (b), and (c) are axial slices that show the fourth ventricle, left white matter, and right vestibule regions, respectively. (d) is a sagittal reformat that shows the CSF surrounding the left facial, vestibular, and cochlear nerves, and (e) and (f) are coronal reformats that show a portion of the right superior semi-circular canal, and the right cochlea, respectively.

| Structure | Mean SNR | Standard Error | 95% Confidence Interval |
|---|---|---|---|
| 4th ventricle | 123.88 | 2.17 | (102.60, 145.16) |
| White matter | 26.49 | 0.44 | (22.21, 30.77) |
| Vestibule | 108.70 | 1.81 | (96.71, 120.70) |
| CSF in nerve region | 93.34 | 1.51 | (83.34, 103.33) |
| SSC | 85.18 | 1.43 | (75.69, 94.66) |
| Cochlea | 92.17 | 1.42 | (82.78, 101.55) |

Table 3.2: SNR of structures examined in the 0.5T bSSFP temporal bone protocol (based on bSSFP acquisitions from n = 21 participants with NEX = 1, no phase cycling).

### 3.3.2 Paranasal Sinus

A representative image acquired with the 0.5T sinus protocol is shown in Figure 3.5, while Figure 3.6 demonstrates the extensive banding present when attempting to image with a similar protocol at 3T (NEX = 1, no phase cycling).

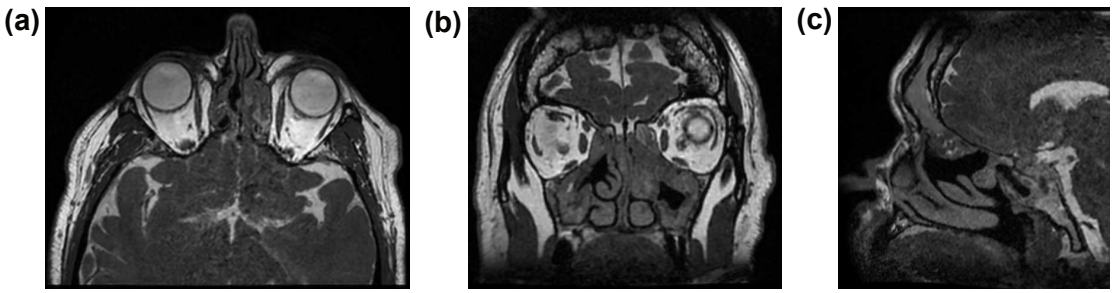


Figure 3.5: A representative image from the 0.5T axial 3D bSSFP paranasal sinus protocol (NEX = 1, no phase cycling). (a) is the native axial plane, (b) is a coronal reformat and (c) is a sagittal reformat.

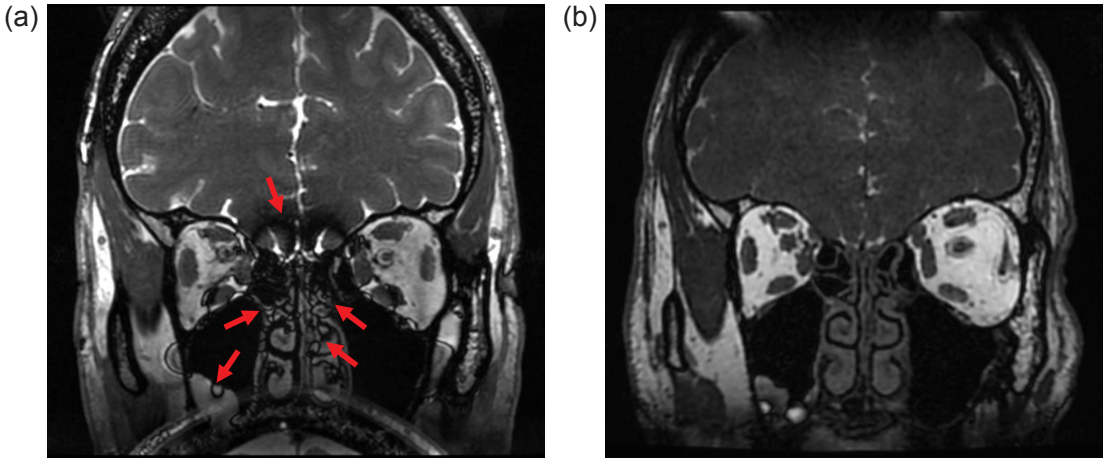


Figure 3.6: A comparison of paranasal sinus acquisitions with similar imaging parameters at (a) 3T (NEX = 1, no phase cycling) and (b) 0.5T (NEX = 1, no phase cycling). The multiple red arrows in (a) point to areas with numerous banding artifacts resulting from off-resonance effects.

Of the 25 participants' images acquired with the 0.5T paranasal sinus protocol (NEX = 1, no phase cycling), 11 (44%) demonstrated banding that intersected the relevant structures, however 4 of these were due to suspected metallic dental work

(Figure 3.7). Therefore, of the 21 images that did not have metallic artifacts intersecting the sinus anatomy, 14 (67%) were free of banding artifact. An example of an acquisition with banding is shown in Figure 3.8.

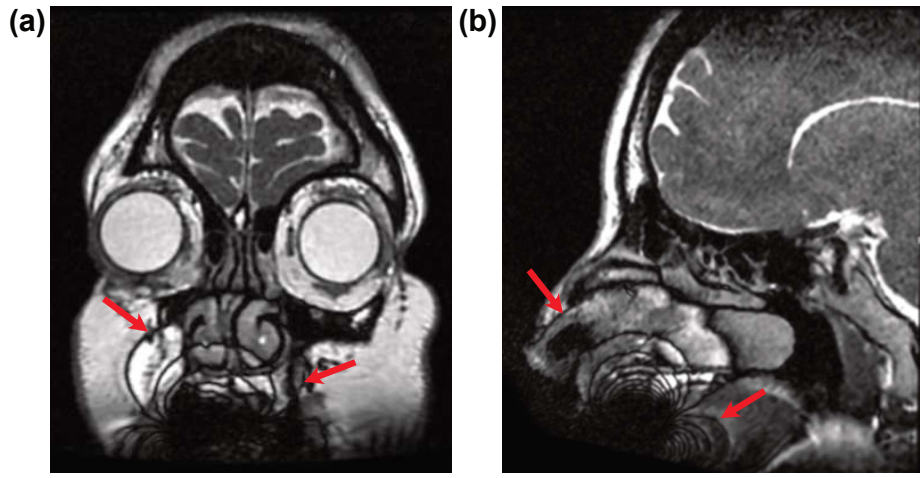


Figure 3.7: Metallic dental work creates dipole field inhomogeneity patterns that lead to bSSFP banding artifacts in 4 participants' 0.5T bSSFP sinus acquisitions (NEX = 1, no phase cycling). One participant's (a) coronal and (b) sagittal reformats are shown here, with red arrows pointing to the outer edges of artifacts resulting from this effect.

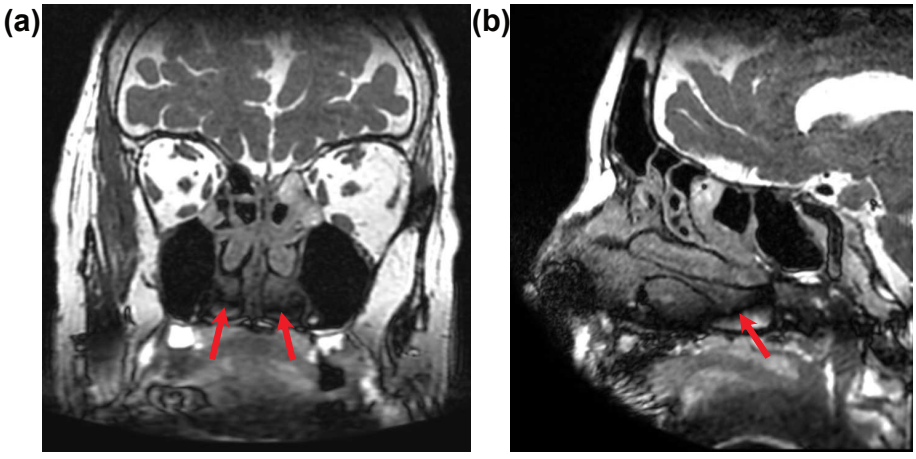


Figure 3.8: Coronal (a) and sagittal (b) reformats of an axial 3D bSSFP acquisition from the sinus protocol developed on the low-field system (NEX = 1, no phase cycling). The red arrows indicate banding artifacts that traverse the ostiomeatal complex and maxillary sinuses in (a) and extend from the jaw to the base of the sphenoid sinus in (b).

The ROIs examined under the 0.5T paranasal sinus protocol are shown in Figure 3.9 and the results of SNR calculations are listed in Table 3.3.

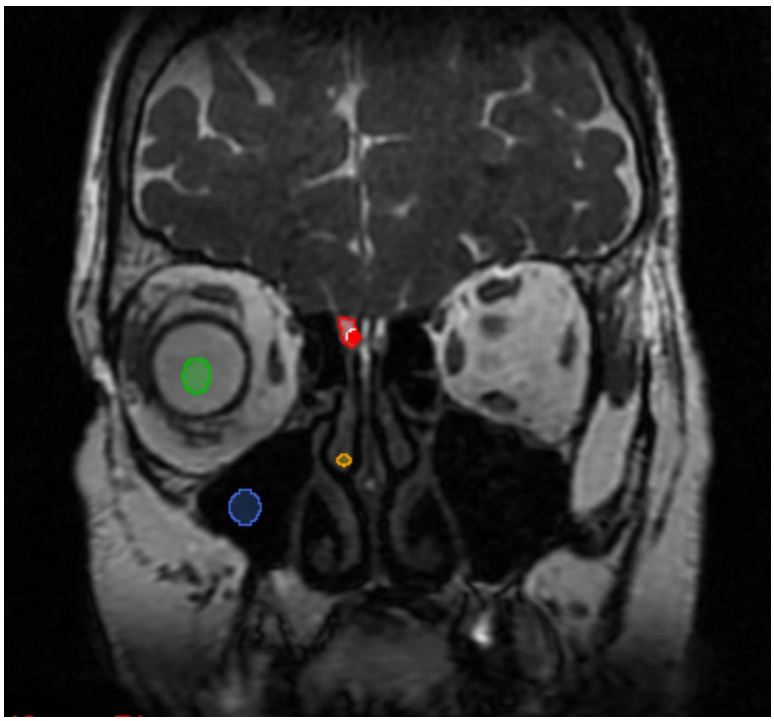

Figure 3.9: The ROIs examined for SNR calculations under the 0.5T paranasal sinus protocol. The red region is the CSF superior to the right cribriform plate, the green region encompasses a portion of the right orbit, the yellow region encompasses some of the mucosa in the right middle turbinate, and the blue region encompasses a portion of the right maxillary sinus.

| Structure | SNR | Standard Error | 95% Confidence Interval |
|---|---|---|---|
| CSF[a] | 81.09 | 1.93 | (60.76, 101.42) |
| Globe | 89.10 | 2.16 | (66.34, 111.86) |
| Maxillary sinus | 3.74 | 0.09 | (2.76, 4.72) |
| Mucosa | 39.89 | 0.98 | (29.59, 50.19) |

Table 3.3: SNR of structures examined in the 0.5T bSSFP paranasal sinus protocol (based on acquisitions from n = 25 participants with NEX = 1, no phase cycling). [a] CSF refers to the CSF superior to the cribriform plate.

## 3.4 Discussion

### 3.4.1 Temporal Bone

More than 50% (81%) of temporal bone images acquired with the 0.5T bSSFP protocol, without phase cycling techniques, were free of banding artifacts traversing relevant structures, allowing us to confirm the first hypothesis. The 3T FIESTA-C (NEX = 2, phase cycling and MIP) acquisitions displayed less artifact tolerance than the bSSFP acquisitions performed at 0.5T, with only 14% being free of banding artifacts traversing relevant structures, thus we can reject hypothesis 2. This is a surprising result, as the FIESTA-C acquisition is commonly used with MIP reconstruction, and is said to "greatly reduce banding artifacts".[60] Given the similar TR values used in both the 0.5T and 3T FIESTA-C protocols, we can infer that the increase in off-resonance effects at 3T was responsible for producing banding in both of the FIESTA-C acquisitions used in the MIP reconstruction. If multiple, large banding artifacts existed at similar locations in each acquisition during phase cycling, these artifacts could persist despite the use of MIP. These results, along with Figure 3.2 demonstrate the advantage of reduced off-resonance effects at low field strengths - a 0.5T bSSFP acquisition of the temporal bone can demonstrate improved artifact tolerance over a similar 3T acquisition that utilizes phase cycling and MIP reconstruction, thereby reducing scan time and eliminating the need for advanced reconstruction techniques.

The SNR of the temporal bone structures examined is well above the minimum required as preferable to radiologists,[76] confirming hypothesis 3 for the temporal bone protocol. While excess SNR suggests that an even higher resolution protocol could be attempted, the required increase in TR would be likely to introduce more banding artifacts in relevant structures. Instead, future studies could explore traditional acceleration techniques, such as GRAPPA,[13] to reduce scan time, while maintaining sufficient SNR to delineate the relevant structures. Further, the fast gradients on the 0.5T system permitted a high readout efficiency, allowing a high SNR acquisition in a shorter scan time than 3T acquisitions.

In summary, the 0.5T bSSFP images of the temporal bone were acquired with scan times less than those of the clinical exams while demonstrating improved artifact tolerance. Additionally, the 0.5T images displayed sufficient SNR at isotropic resolutions

equivalent to clinical acquisitions. These results support the case for future use of the 0.5T system for clinical temporal bone exams.

### 3.4.2 Paranasal Sinus

The artifact tolerance of the paranasal sinus protocol was shown to be sufficient to confirm hypothesis 1, i.e. $> 50\%$ of the images were free of artifacts traversing relevant structures. Even more artifact tolerance could be prioritized in future protocol modifications; a decrease in resolution would permit an even shorter TR, making the acquisition even more robust to off-resonance effects. Future studies, as well as clinicians' opinions, are required to determine whether to prioritize resolution or artifact suppression in future versions of the 0.5T paranasal sinus protocol. The attempt to acquire an image with similar acquisition parameters on a 3T system (Figure 3.6) further demonstrates the advantages of imaging with a low-field system with high-performance gradients - to acquire the same resolution and FOV, a TR/TE of 4.2ms / 1.6ms is required at a RBW of 73 kHz to image at 3T (compared to 3.5ms / 1.6ms and 150 kHz at 0.5T), due primarily to SAR limits, as well as the lower gradient slew rate (140 T/m/s at 3T compared to 200 T/m/s at 0.5T). The necessity of a long TR creates banding that renders the images non-diagnostic.

When imaging at 0.5T, the high SNR efficiency of the bSSFP sequence allowed a high receive bandwidth to be tolerated for banding artifact suppression via shortened TR, while maintaining SNR sufficient for radiologist visualization. The high SNR values ($> 39$) found in structures other than the air-filled maxillary sinus (Table 3.3), permits the delineation of structures and allows us to confirm hypothesis 3 for the paranasal sinus protocol. These results suggest that SNR could be sacrificed such that even higher resolution protocols could be explored in the future. Doing so would require employing higher receive bandwidths with stronger, faster slew rate gradient settings. Notably, the 0.5T system used has an upper slew rate setting of 400 T/m/s without producing evident peripheral nerve stimulation from testing at our facility, however gradient heating must be taken into account when high slew rates are used.

While CT is the current standard of imaging for paranasal sinuses, our health center performs CT acquisitions with a slice thickness of 3mm. Our MRI protocol offers preferable, sub-millimeter resolution in the coronal and sagittal planes, as well

as superior soft-tissue differentiation, and is free of potentially harmful radiation. Furthermore, the parameters used in our sinus acquisition resulted in a scan time of 1 min 52s, which competes with CT when practical aspects of clinical work flow are considered. Combined with the high SNR and resolution of the acquired images, this work demonstrates the potential of MRI for high throughput sinus pathology evaluation.

Overall, the quantitative evaluation of the 0.5T temporal bone and paranasal sinus protocols performed here shows that the decrease in off-resonance effects at 0.5T, and the ability to use a short TR, are beneficial in reducing the banding artifacts that are expected when imaging regions with susceptibility differences. Moreover, imaging with a head-only system with strong, fast gradients permits artifact-tolerant acquisitions without sacrificing SNR or resolution. We suspect that post-processing noise suppression from proprietary reconstruction techniques is, in part, responsible for the resulting SNR values. It is reasonable to hypothesize that the remaining contributions to SNR resulted from the use of the conformable 16-channel head-coil (compared to the rigid head-coil on the 3T system), while the additional elements of the modern receive chain no doubt aided in the maintenance of sufficient SNR.

### 3.4.3 Limitations

Although it does not represent a true limitation of the study, a comprehensive comparison of artifact tolerance and SNR for 0.5T and clinical paranasal sinus acquisitions was not feasible, given that clinical MRI of the paranasal sinuses is not performed at our institution. Importantly, while the metrics examined here are informative, the true value of an MR acquisition lies in a radiologist's ability to use the image to answer a clinical question. To that end, a study that investigates radiologists' ability to visualize structures of the temporal bone is presented in the following Chapter.

## 3.5 Conclusions

The results of this study support the confirmation of Hypothesis 1, i.e. >50% of the images in each protocol were free of artifact traversing relevant structures. Hypothesis

2 was rejected in that the 0.5T temporal bone protocol (NEX = 1, no phase cycling) produced more images that were free of banding artifacts traversing relevant structures than the clinical 3T FIESTA-C protocol (NEX = 2, phase cycling and MIP). The SNR from the 0.5T temporal bone and paranasal sinus protocols was shown to be greater than that preferred for visualization by radiologists, allowing us to confirm Hypothesis 3.

# Chapter 4

# Low-field vs. conventional field balanced steady-state free precession imaging of the temporal bone: radiologist rating of anatomical visualization

## 4.1 Study Objectives and Hypotheses

The objective of this study is to evaluate the advantages and disadvantages of our 0.5T temporal bone bSSFP protocol in a clinical context. To achieve this objective, we will examine radiologists' ability to visualize structures of the inner ear and IAC in bSSFP images acquired at 0.5T (NEX = 1, no phase cycling) in comparison to the clinical FIESTA-C images acquired at 3T (NEX = 2, phase cycling and MIP).

While full-reference image quality metrics (IQM), such as root mean square error (RMSE) and structural similarity index (SSIM), are commonly used to facilitate quantitative comparison of images, none are suited for application across field strengths. SSIM was considered for use in this study, however, it is a measure of similarity in luminance, contrast, and structural content,[78] and it can be expected that image luminance and contrast will vary across different field strength acquisitions. In addition, it has been shown that various IQMs do not yield strong correlations to radiologists' evaluation of image quality. Therefore, this study makes use of a task-based measurement, instead of full-reference IQMs, to compare image quality.[78]

Clinical MRI of the paranasal sinuses is not performed at our institution, therefore a comparison analogous to that being done with temporal bone protocols cannot be performed.

The results of the previous study showed that the protocol exhibits better artifact tolerance compared to 3T FIESTA-C acquisitions, and sufficient SNR for visualization by radiologists, with isotropic resolution identical to the clinical acquisitions. We therefore hypothesize that:

1. Radiologists' ratings of their ability to visualize the SSC, vestibule, facial nerve,

cochlear nerve, and cochlea will be significantly higher for images acquired at 0.5T than at conventional field strengths. Consequently, we also hypothesize that

2. the combination of all ratings at each field strength will result in a rating that demonstrates a statistically significant improvement in visualization of structures in images acquired at 0.5T.

## 4.2 Methods

### 4.2.1 Image Acquisition

An objective inclusion criteria of participant isocenter being $< 60mm$ from the imaging isocenter was used to select image sets to be included in this study. This criterion was chosen as this position was determined to be the location above which concomitant gradient artifacts began to obscure the relevant anatomical structures (see Section 2.4.2). In addition, we required that participants' clinical acquisitions were performed at a field strength of 3T.

A subset of 19 images fit this criteria, however 15 were included in the study - 3 participants had large vestibulocochlear schwannomas that obscured relevant anatomy, and 1 participant's images were excluded due to an error in the 0.5T pre-scan calibration that rendered the image non-diagnostic. 7 of the participants identified as male ($52 \pm 13$ years), and 8 identified as female ($47 \pm 15$ years), with a mean age of 49 years and standard deviation of 14 years.

Of the 15 included data sets, 6 of the 0.5T acquisitions were acquired with the original temporal bone protocol using TR = 6.7ms, and a maximum gradient strength of 60 mT/m. The remaining 9 were acquired with the modified protocol with TR = 7.0 ms, and a maximum gradient strength 30 mT/m. All other parameters are shown in Table 2.1). Images from 14 of the participants were included in the previous study.

SOC imaging consisted of FIESTA-C acquisitions (NEX = 2, and phase cycling with MIP) at 3T (GE Discovery MR750), using the same parameters shown in Table 3.1. The FIESTA-C sequence consisted of one acquisition with an RF phase pulse that alternated between 0° and 180°, and one in which the phase of the RF pulse remained constant.

### 4.2.2 Image Analysis

All images were anonymized and shown to two board-certified neuroradiologists (each with $>$ 5 years of sub-speciality experience), and the order in which the 3T and 0.5T images for each participant were provided was randomized. The radiologists independently rated their ability to visualize 5 anatomical structures relevant to clinical temporal bone MRI exams: the superior semicircular canal (SSC), the vestibule, the facial nerve, the cochlear nerve, and the cochlea. Ratings were reported using a 5-point Likert scale ranging from 1 (very poor) to 5 (excellent). Scoring was separately recorded for left and right ears. During analysis, left and right ratings were concatenated such that there was one set of ratings per anatomical structure for each rater.

To quantify inter-rater reliability, Gwet's AC2 (with ordinal weights) was calculated for each structure at each field strength with the AgreeStat360 application.[79] There is strong evidence that low kappa values can occur when utilizing Cohen's kappa on ordinal data with high agreement[80–82] and Gwet's AC2 has been shown to overcome this paradox. On the advice of our group's statistician, results from weighted Gwet's AC2 tests are presented in the following section with their associated 95% confidence intervals.

Wilcoxon signed rank tests were used to compare ratings at each field strength cohort for each structure. In addition, an overall comparison was made in which the combined scores at 0.5T were compared to those at 3T. Python 3.6 and the Statistical Functions module were used to calculate $P$ values. A result of $P < 0.05$ was considered significant. Since radiologists' ability to visualize a given structure was assumed to be independent from their ability to visualize another, it was unnecessary to apply a correction factor for multiple comparisons.

## 4.3 Results

Representative images acquired at each field strength are shown in Figure 4.1, and demonstration of the structures being examined by radiologists is shown in Figure 4.2.

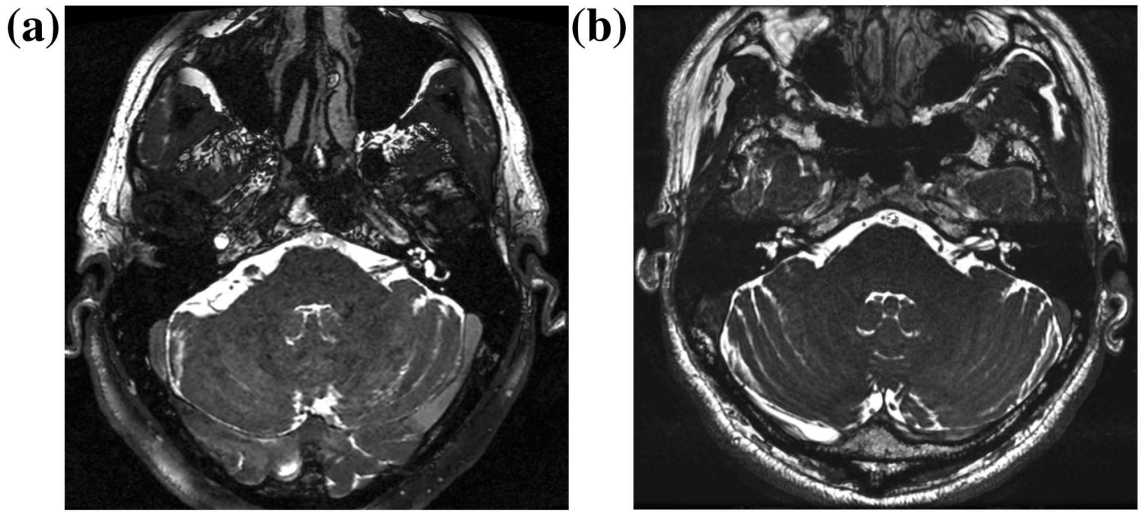


Figure 4.1: Representative axial bSSFP MR images of the temporal bone from two participants. Panel (a) displays an example slice from axial bSSFP acquisitions on the 0.5T system (NEX = 1, TR/TE 7.0 ms/3.3 ms, scan time 4 min 19 s. All other imaging parameters are shown in Table 2.1). Panel (b) displays an example FIESTA-C acquisition from the 3T system (NEX = 2, TR/TE 5.5 ms/2.2 ms, number of slices 136, scan time 5 min 20 s, phase cycling and MIP. All other imaging parameters as shown in Table 3.1).

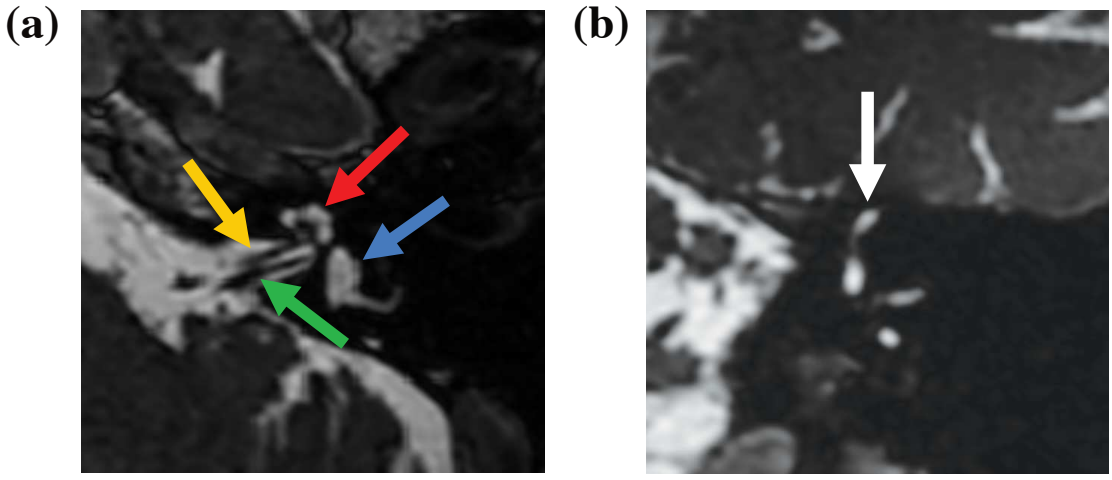


Figure 4.2: A depiction of the structures being rated. Images were acquired at 0.5 T (TR/TE 7.0 ms/3.3 ms, maximum gradient strength 30 mT/m, scan time 4 min 19 s. All other imaging parameters are shown in Table 2.1). Image (a) is an axial slice depicting the left cochlea (red arrow), vestibule (blue arrow), cochlear nerve (yellow arrow) and facial nerve (green arrow), and image (b) is a coronal reformat where the white arrow points to the SSC.

The results of the Gwet's AC2 test for inter-rater reliability are shown in Table 4.1. According to the scale defined by Landis and Koch,[83] the results demonstrate moderate to substantial agreement ($0.41 < AC2 < 0.80$) between raters' Likert scores for all structures except the cochlea which showed fair agreement ($0.21 < AC2 < 0.40$).

| | Gwet's AC2 Value (Ordinal Weights) | 95% Confidence Interval |
|---|---|---|
| **Conventional Field Strengths** | | |
| Overall | 0.8 | (0.71, 0.80) |
| SSC | 0.8 | (0.66, 0.87) |
| Vestibule | 0.8 | (0.71, 0.87) |
| Facial nerve | 0.9 | (0.79, 0.93) |
| Cochlear nerve | 0.9 | (0.86, 0.94) |
| Cochlea | 0.5 | (0.45, 0.60) |
| **0.5T** | | |
| Overall | 0.9 | (0.85, 0.9) |
| SSC | 0.9 | (0.88, 0.97) |
| Vestibule | 0.9 | (0.83, 0.90) |
| Facial nerve | 0.9 | (0.88, 0.96) |
| Cochlear nerve | 0.9 | (0.85, 0.94) |
| Cochlea | 0.8 | (0.70, 0.86) |

Table 4.1: The results of Gwet's AC2 calculations for inter-rater reliability and their 95% confidence intervals.

The ratings are summarized by boxplots presented in Figure 4.3. Raters' Likert scores, overall, do not differ significantly when rating their ability to visualize structures of the temporal bone on images acquired 0.5T (NEX = 1, no phase cycling) or 3T (NEX = 2, phase cycling with MIP), resulting in a $P$ value of 0.20, however, there are differences when specific structures are being considered. Raters prefer the images acquired at 0.5T for visualizing the SSC ($P < 0.001$) and the vestibule ($P = 0.006$), whereas 3T acquisitions were preferred for visualizing the facial and cochlear nerves ($P = 0.003$ and $P = 0.004$, respectively). No significant difference was found between ratings for each field strength cohort when visualizing the cochlea ($P = 0.70$).

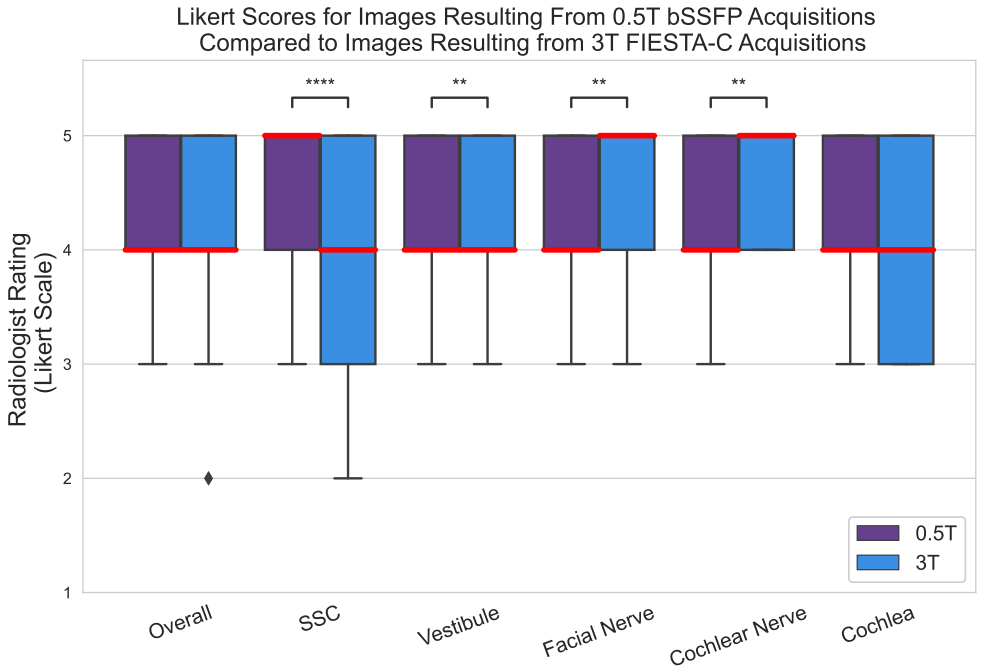


Figure 4.3: Ratings for visualization of structures at 0.5T and 3T. The bold, red lines indicate median values. * indicates $0.01 < P \leq 0.05$, ** indicates $0.001 < P \leq 0.01$, and **** indicates $P \leq 0.0001$. The lack of whiskers indicating the upper quartile range reveals that the ratings were skewed towards the upper limit of the Likert scale.

## 4.4 Discussion

Interestingly, we were not able to confirm or reject Hypothesis 1 uniformly. There was no statistically significant difference in overall ratings for images acquired at 0.5T (NEX = 2, no phase cycling) and 3T (NEX =1, phase cycling with MIP), i.e. we rejected Hypothesis 2 with $P = 0.2$, however, significant differences were found for ratings of four of the five structures examined.

The results allowed us to confirm Hypothesis 1 for the SSC and vestibule, as raters preferred the 0.5T acquisitions for visualizing these structures ($P < 0.001$ and $P = 0.006$, respectively). Subsequent discussions with raters revealed that these preferences were due to a lower prevalence of banding; some of the FIESTA-C images acquired at 3T demonstrated fine banding artifacts that traversed the region surrounding the vestibule (Figure 4.4), thereby making dehiscence of the nearby SSC difficult to confirm. This demonstrates the advantages, in the context of a clinical temporal bone exam, of decreased off-resonance effects when imaging a low field

strengths - it is possible to acquire artifact-tolerant bSSFP images at 0.5T, without the need for phase cycling techniques, permitting sufficient visualization in scan times less than that required when imaging at 3T.

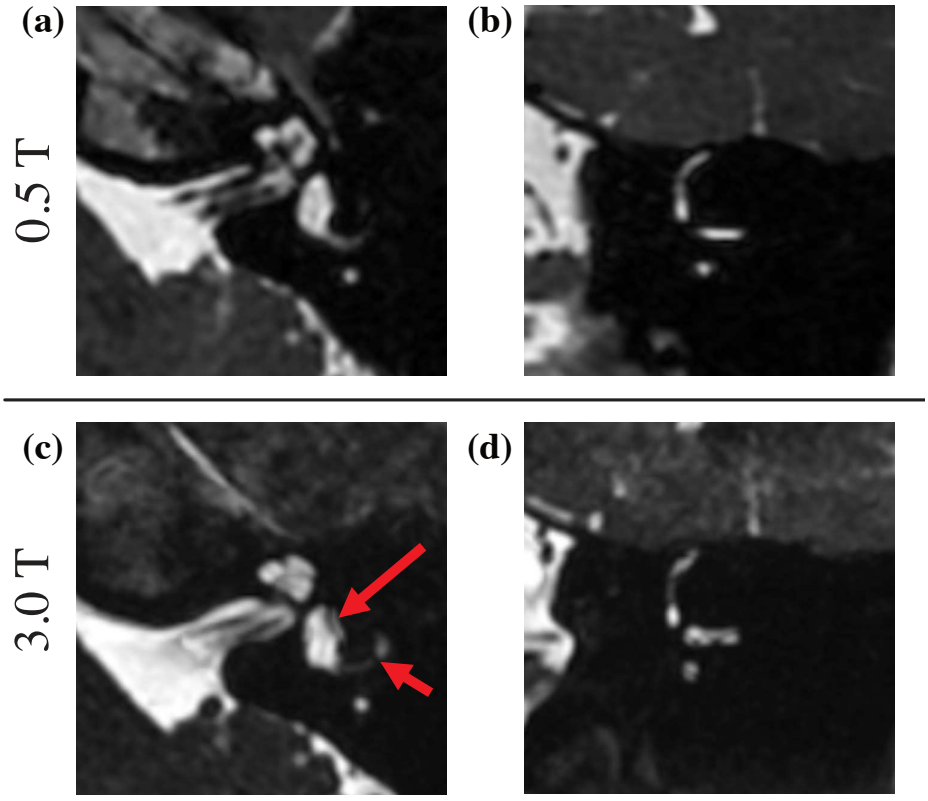


Figure 4.4: Example in which the 0.5T acquisition (top panel) was preferred over the 3T acquisition (lower panel) for visualizing the vestibule and SSC by both raters. All panels display coronal reformats of axial 3D bSSFP acquisitions. 0.5T data were acquired with NEX = 1, no phase cycling, TR/TE 6.7 ms / 3.3 ms, and a scan time of 4min. 3T FIESTA-C data was acquired with NEX=2, TR/TE = 5.5/2.2ms, 136 slices, scan time 5min 20s, and is presented as a MIP of the phase-cycled acquisitions. The red arrows in (c) indicate examples of banding artifacts that are seen in the vestibule in some 3T FIESTA-C acquisitions.

Hypothesis 1 was rejected for the facial and cochlear nerves, as raters preferred 3T FIESTA-C acquisitions when visualizing these structures ($P = 0.003$ and $P = 0.004$, respectively). Raters expressed that these preferences were due to the nerves appearing "nodular" when visualized at 0.5T, making it difficult to rule out tiny tumours. The difference in appearance between field strength cohorts is demonstrated in Figure 4.5. A possible decrease in SNR in images acquired at 0.5T, as well as

variations in reconstruction algorithms across vendors may be responsible for these preferences.

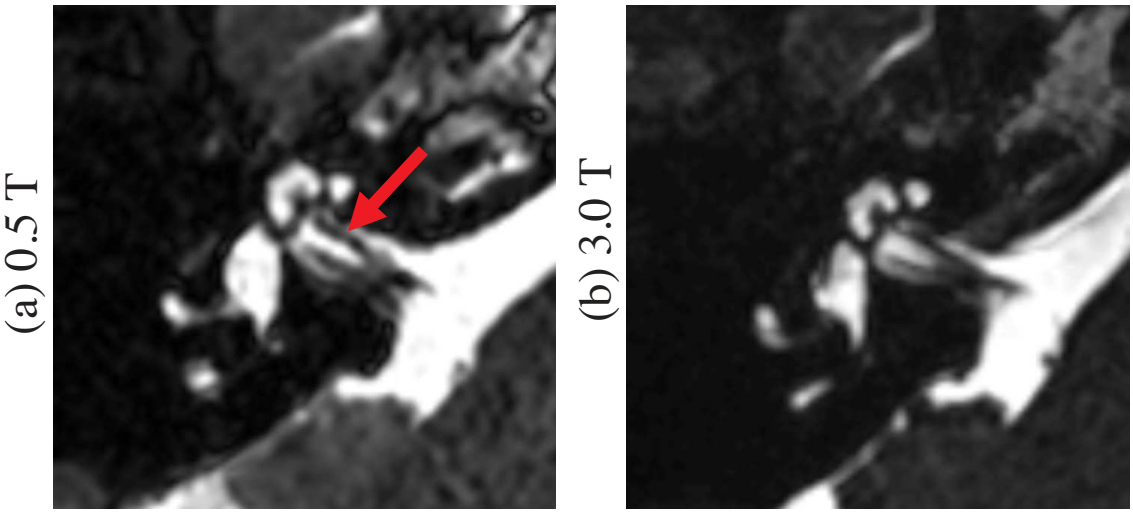


Figure 4.5: An example in which the (b) 3T FIESTA-C acquisition was preferred over the (a) 0.5T acquisition for visualization of the cochlear and facial nerves. The red arrow in (a) indicates the "nodular" feature that raters disliked. 0.5T data were acquired with NEX = 1, TR/TE 6.7 ms/3.3 ms, and a scan time of 4min. 3T data were acquired with NEX = 2, TR/TE = 5.5 ms/2.2 ms, and a scan time of 5min 26s, and is presented as a MIP of the phase-cycled acquisitions.

While CT is the current standard to diagnose SSC dehisence, MRI is required to rule out tumours of the inner ear and schwannomas along the nerves in the inner auditory canal.[50] Our results do not demonstrate improved visualization of nerves at 0.5T, and additional studies should focus on the accuracy of tumour diagnosis, and other areas of diagnostic utility, of low-field bSSFP imaging of the temporal bone.

Hypothesis 2 is rejected since the 0.5T images were not rated as significantly higher than the images acquired at 3T when all ratings were combined to perform an overall comparison. This, however, is still an important finding - bSSFP images acquired at 0.5T, without phase cycling techniques, are not rated significantly different from images acquired at 3T with phase cycling and MIP. Performing temporal bone imaging at 0.5T therefore has the potential to reduce scan times and thereby reduce the overall burden on the healthcare system.

### 4.4.1 Limitations

The primary limitation of this study was the inability to truly blind raters to the field strength used to acquire a given image; this is an unfortunate and unavoidable result of radiologist training and experience that could have led to unconscious bias during rating. Secondly, the majority of participants underwent 0.5T scans immediately following their SOC exam, and it is possible that participants became uncomfortable over time, causing motion-induced artifacts in the 0.5T acquisitions. This could have had the effect of causing lowered ratings of the 0.5T images. Although a comparison of paranasal sinus images acquired at 0.5T and 3T was not feasible, a parallel study is being performed by other members of the group, which compares radiologists' ability to answer clinical questions when using images from our 0.5T paranasal sinus protocol to their ability when using participants' clinical CT images.

## 4.5 Conclusions

Hypothesis 1 was confirmed for the SSC ($P < 0.001$) and the vestibule ($P = 0.006$), as raters preferred the 0.5T acquisitions for visualizing these structures. Hypothesis 1 was rejected for the facial and cochlear nerves, as raters preferred the 3T FIESTA-C acquisitions for visualizing these structures ($P = 0.003$, and $P = 0.004$, respectively). Hypothesis 1 was also rejected for the cochlea, since there was no significant difference between ratings ($P = 0.7$). The overall ratings did not demonstrate significant improvements for visualization at 0.5T ($P = 0.2$), thus Hypothesis 2 is rejected.

# Chapter 5

# Super resolution allows for the benefits of low resolution balanced steady-state free precession imaging without degrading image quality.

## 5.1 Study Objectives and Hypotheses

As described, compressed sensing is commonly used to decrease image acquisition times, thereby reducing patient motion. However, in addition to reduced scan times, patient motion, and patient heating, low resolution (LR) bSSFP acquisitions can benefit from the use of a shorter TR, enabling improved artifact tolerance. Importantly, however, high resolution (HR) is a priority for temporal bone imaging to ensure sufficient visualization of the fine structures of the inner ear.

The overall objective of this study is, therefore, to acquire LR images of the temporal bone and to return them to HR with a publicly available super resolution (SR) machine learning model, without significant degradation.

Many studies that implement super resolution pipelines synthesize LR images from HR acquisitions (e.g.[26,84]). In doing so, there are paired true HR and super resolution HR images that can be used to quantify the performance of the model, by, for example, RMSE and/or SSIM.[85] It is relevant to note that since our approach involves separate HR and LR acquisitions, accurate co-registration of the true HR and super resolution HR images would be required to utilize a pixel-wise comparison such as RMSE. Further, while SSIM calculations were attempted, it was found that the results were dominated by imperfections in co-registration, and that expected differences in image characteristics rendered the results of SSIM calculations uninformative. Instead, this study employs two board-certified neuroradiologists to rate their ability to visualize temporal bone structures in images from the original temporal bone protocol ("original resolution", or OR) as well as LR acquisitions, both brought to HR with standard interpolation methods, and also with a super resolution

pipeline. Analysis will compare ratings of the OR images brought to HR with standard interpolation methods to LR images brought to HR with the super resolution pipeline. In addition, the effect of each method applied will be examined by comparing ratings between HR images resulting from LR acquisitions (with interpolation and super resolution), and a similar comparison will be made between ratings of the images resulting from OR acquisitions.

We hypothesize that:

1. LR images brought to HR with super resolution methods will not be rated significantly different, or will be rated significantly higher, than HR images brought to clinical resolution with standard interpolation techniques.

2. LR images brought to HR with super resolution methods will be rated significantly higher than LR images brought to HR with standard interpolation methods.

3. OR images brought to HR with super resolution methods will be rated significantly higher than OR images brought to HR with standard interpolation methods.

## 5.2 Methods

### 5.2.1 Image Acquisition

Images were acquired from 10 healthy volunteers. 6 participants identified as male (mean age 38 $\pm$ 11 years) and 4 identified as female (mean age 28 $\pm$ 1 years). The total participant cohort had a mean age of 35.6 years $\pm$ 11. Three axial, 3D bSSFP images were acquired from each patient: two images were acquired with the original 0.5T temporal bone protocol, which used a TR of 7.0ms, and an isotropic resolution of 0.6 mm. These images are therefore referred to as the "original resolution", OR, acquisitions. One OR acquisition was interpolated onto a 0.3 mm image matrix with matrix size ($600 \times 600 \times 328$) (as described previously, e.g. Table 2.1), and one was acquired without interpolation, generating an image with matrix size ($300 \times 300 \times 164$). One LR image was acquired with a TR of 6.3 ms, and an isotropic resolution of 0.7 mm without interpolation performed, generating an image with matrix size ($260 \times$

| Imaging Parameter | $OR_{Int.}$ | OR | LR |
|---|---|---|---|
| Field of view (cm) | 18 | 18 | 18 |
| Acquired isotropic resolution (mm) | 0.6 | 0.6 | 0.7 |
| Acquisition matrix | (300 × 300) | (300 × 300) | (260 × 260) |
| Number of slices | 164 | 164 | 140 |
| Receive Bandwidth (kHz) | 70 | 70 | 70 |
| Repetition time (ms) | 7.0 | 7.0 | 6.3 |
| Echo time (ms) | 3.4 | 3.4 | 3.1 |
| Flip angle (°) | 60 | 60 | 60 |
| Interpolated isotropic resolution (mm) | 0.3 | 0.6 | 0.7 |
| In-plane image matrix size | (600 × 600) | (300 × 300) | (260 × 260) |
| Number of image slices | 328 | 164 | 140 |
| Scan time | 4min. 19s | 4min. 19s | 3min. 11s |

Table 5.1: Imaging parameters for high resolution (HR) acquisitions with and without interpolation, and low resolution (LR) acquisitions without interpolation

$260 \times 140$). The imaging parameters used for these 3 protocols are shown in Table 5.1. Notably, the scan time of the 0.7mm protocol was 3 minutes, 11 seconds, which is a decrease in scan time of 1 minute 8 seconds compared to the 0.6mm protocol.

### 5.2.2 Data Set Preparation

This section describes the process of generating an experimental data set for each participant. For reference, a summary is provided at the end of this section in Figure 5.7.

#### Interpolation

As demonstrated in Table 5.1, one of the OR protocols had interpolation to 0.3mm performed as part of the standard reconstruction performed by the system. Images obtained with this protocol were included as one of the experimental data sets and are henceforth referred to as the $OR_{Int.}$ cohort.

Bicubic interpolation of the LR images from $(260 \times 260 \times 140)$ to $(600 \times 600 \times 328)$ was performed with the Open Source Computer Vision Library (OpenCV) for Python 3.6, generating an image matrix with 3D isotropic resolution of 0.3mm. The resulting images are henceforth referred to as the $LR_{Int.}$ cohort.

## Super Resolution

### Network Architecture and Training

The SR model we employed was the "Image Super-Resolution (ISR)" model provided by Cardinale.[86] The model architecture is an implementation of the residual dense network (dense ResNet) created by Zhang et al,[67] shown in Figure 5.1. The RDN consists of 20 residual dense blocks (RDBs), each with 6 convolutional layers that extract 64 feature maps from their input. The convolutional layers outside the RDBs also produce 64 feature maps.

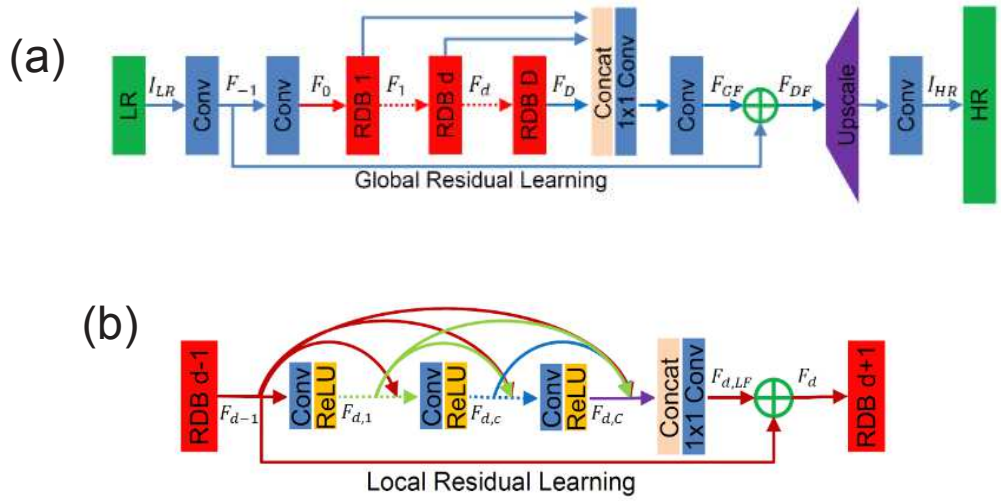


Figure 5.1: (a) The architecture of the residual dense network used in ISR, developed by Zhang et al. "Global residual learning" is achieved by summing the shallow feature maps, $F_{-1}$ with the global feature maps, $F_{GF}$ to produce the dense feature maps, $F_{DF}$. The upscaling was performed by using the Efficient Sub Pixel Convolutional Neural Network (ESPCNN) designed by Shi et al.[87] Each of the residual dense blocks (RDBs) in (a) performs "local residual learning" (b) via summation of the feature maps from the previous layer, $F_{d-1}$ with the local feature maps, $F_{d,LF}$, which are created by the current residual dense block. The figure is generated from those presented by Zhang et al 2018.[67]

Multiple sets of pre-trained weights for the model were provided. One such set resulted from training the model on the DIV2K dataset,[88] while optimizing the mean squared error (MSE) and validating with the peak signal-to-noise ratio (PSNR). As this training is described as "PSNR-driven", the resulting weights are henceforth referred to as $W_{PSNR}$.

For a second set of weights, the RDN was trained by minimizing a combination of MSE, and perceptual loss (using VGG19). In addition, the RDN was treated as the generating component of a GAN such that the prediction produced by the model was fed to a discriminator, where the discriminator architecture was based on that designed by Ledig et al.,[24] shown in Figure 5.2. The overall loss used to produce the second set of weights was therefore calculated as a weighted sum of MSE, perceptual loss, and the binary cross-entropy resulting from the GAN component. Model output was validated using PSNR. This architecture was trained, in part, on input images from the DIV2K data set, compressed by 50%. Since the model then learned to reduce compression artifacts, the resulting weights were referred to as the "noise-cancelling" weights by the authors, and we will refer to them as $W_{NC}$ here. Further details on network architecture and training can be found at Cardinale's GitHub repository.[86]

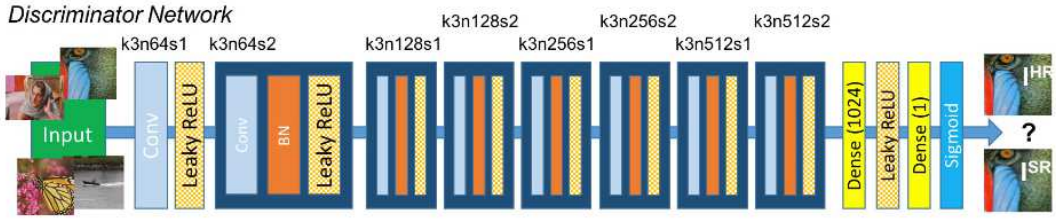


Figure 5.2: The architecture of the discriminator used to produce $W_{NC}$ in the ISR model. Here, k is the dimension of the square convolution kernel, n is the number of feature maps extracted, and s is the stride used in each step of the convolutional kernel. The figure is reproduced from Ledig et al. 2017.[24]

### Pre- and Post-processing of Images

ISR was designed to operate on square, 2D, RGB images of UINT8 (8-bit unsigned integer) datatype, and as such, a number of pre- and post-processing steps were required to use the network for our application. These steps were completed using Python 3.6 with a number of required image processing packages. An overview of the steps performed for one image volume is provided here. For clarity, the dimensions of the output at each step in the SR pipeline are shown for LR and OR (without interpolation) input in Table 5.2.

Images from the 0.5T system are automatically saved in DICOM format, with INT16 pixel arrays. To begin, the DICOM image volume was loaded into an image array with dimensions $(x, y, z)$, corresponding to the anterior-posterior, left-right, and superior-inferior directions, respectively. Pixel values were re-scaled to the interval $[0, 255]$ for conversion to UINT8.

The array was zero-padded along the slice dimension ($z$) to match the $x$ and $y$ dimensions, thereby generating a cubic image volume. The image volume was then converted to UINT8 data type. To simulate a 3-channel RGB image volume, the volume was repeated along a fourth array dimension (i.e. $(x, y, z, 3)$). Each 2D image slice along the superior-inferior, anterior-posterior, and left-right directions were saved as individual Portable Network Graphics (.png) files creating 2D square, RGB, images representing axial, coronal, and sagittal slices, respectively, to be fed to the ISR network. The series of images along each direction were then individually sent through the ISR model such that 3 "SR" image volumes were produced.

The zero-padding along the $z$-direction was propagated through the ISR network, therefore the first post-processing step was to crop the excess rows in the image volumes resulting from coronal and sagittal image inputs. Next, since the number of slices fed to the ISR network is not altered, this dimension is not doubled by the network. To produce an image volume that is doubled in all dimensions, each SR volume underwent bicubic interpolation along this direction. A complete estimation of a 3D SR image was then obtained by taking the average of the three resulting image volumes, and the final image array was converted back to gray scale. Converting the image volumes to UINT8 datatype altered the pixel values compared to the original INT16 input, therefore the final SR image volume was converted back to UINT16 datatype and normalized to the maximum of the OR image volume, in an attempt to create comparable contrast.

Notably, the ISR model only permits increases in resolution by factors of 2. Each LR image was therefore brought from $(260 \times 260 \times 140)$ to $(520 \times 520 \times 280)$ with the SR pipeline. These images therefore underwent additional bicubic interpolation to $(600 \times 600 \times 328)$ to match the dimensions of the other images in each data set. LR images brought to clinical resolutions with the SR pipeline are therefore referred to as the $LR_{SrInt.}$ data set. The OR acquisitions did not require additional interpolation

and were brought to HR with all steps discussed previously, creating the $OR_{Sr}$ data set.

| | Image Array Size $(x, y, z)$ | |
|---|---|---|
| | LR Input | OR Input |
| | $(260, 260, 140)$ | $(300, 300, 164)$ |
| 1. Pad image volume along slice dimension | $(260, 260, \mathit{260})$ | (300, 300, *300*) |
| 2. Convert to UINT8 | (260, 260, 260) | (300, 300, 300) |
| 3. Simulate RGB image | (260, 260, 260, *3*) | (300, 300, 300, *3*) |
| 4. Run ISR on: | | |
| Axial slices | (*520*, *520*, 260, 3) | (*600*, *600*, 300, 3) |
| Coronal slices | (260, *520*, *520*, 3) | (300, *600*, *600*, 3) |
| Sagittal slices | (*520*, 260, *520*, 3) | (*600*, 300, *600*, 3) |
| 5. Crop excess padding along slice dimension | | |
| Axial slices | (520, 520, 260, 3) | (600, 600, 300, 3) |
| Coronal slices | (260, 520, *280*, 3) | (300, 600, *328*, 3) |
| Sagittal slices | (520, 260, *280*, 3) | (600, 300, *328*, 3) |
| 6. Interpolate volumes along direction that images were fed to ISR | | |
| Axial slices | (520, 520, 280, 3) | (600, 600, 328,3) |
| Coronal slices | (*520*, 520, 280, 3) | (*600*, 600, 328, 3) |
| Sagittal slices | (520, *520*, 280, 3) | (600, *600*, 328, 3) |
| 7. Calculate mean image volume | (520, 520, 280, 3) | (600, 600, 328, 3) |
| 8. Convert to gray scale | (520, 520, 280) | (600, 600, 328) |
| 9. Convert to INT16 | (520, 520, 280) | (600, 600, 328) |
| 10. Interpolate to same dimensions as $OR_{Int.}$, $LR_{Int.}$ and $OR_{Sr}$ | (*600*, *600*, *328*) | N/A |

Table 5.2: Pre- and post-processing steps for creation of super resolution image volumes. Dimensions that have been changed from the previous processing step are denoted in italics.

Representative axial slices resulting from LR input to the SR pipeline, with the

use of $W_{PSNR}$ and $W_{NC}$ are shown in Figure 5.3. As is clear, both sets produced undesirable outcomes; the use of $W_{PSNR}$ resulted in images with point-like high signal artifacts, while the use of $W_{NC}$ produced images that appear over-smoothed and unrealistic. Further, both sets of weights resulted in images with reduced contrast compared to the LR acquisition.

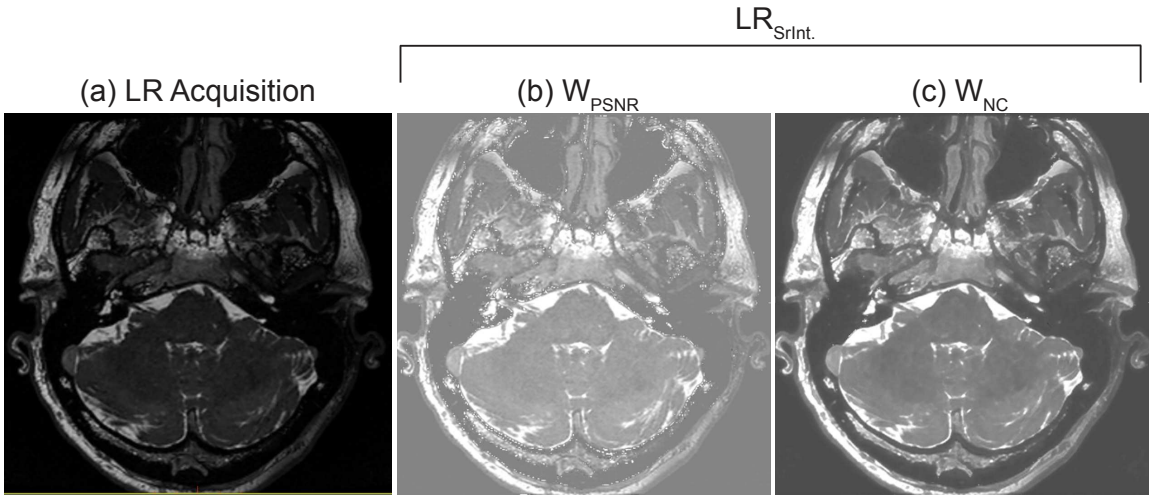


Figure 5.3: Axial slices resulting from the (a) LR acquisition sent through the SR pipeline using (b) $W_{PSNR}$ and (c) $W_{NC}$.

### Interpolating Network Weights

In an attempt to improve on the images generated by the SR pipeline, we utilized the method of Wang et al.[89] to create a new set of model weights resulting from interpolation between $W_{PSNR}$ and $W_{NC}$, i.e.

$$W_{Interp.} = (1 - \lambda)W_{PSNR} + \lambda W_{NC}. \tag{5.1}$$

To determine the value of $\lambda$, images of the small ACR phantom[90] were acquired with the $OR_{Int.}$ and LR sequences shown in Table 5.1. The LR image volume was then brought to 0.3mm isotropic resolution with the SR pipeline (Table 5.2), using $W_{PSNR}$ (i.e. $W_{Interp.}$ with $\lambda = 0$), $W_{NC}$ (i.e. $W_{Interp.}$ with $\lambda = 1.0$), and $W_{Interp.}$ with $\lambda = 0.5$.

To objectively quantify the performance of the SR pipeline with each value of $\lambda$, we began by examining the resolution of each outcome. To do so, we used an

axial image slice with the resolution insert visible. The insert consists of three hole arrays, with hole diameters of 0.9mm, 0.8mm, and 0.7mm when viewed from left to right. The profile of pixel intensities across the top horizontal row of all three upper left hole arrays, and the profile of pixel intensities down the right-most vertical column of the 0.8mm lower hole array were obtained using MATLAB_R2018a for each $LR_{SrInt.}$ result. In addition, the $LR$ acquisition was examined to estimate whether the SR pipeline improved resolution, and the $OR_{Int.}$ image was also examined as it demonstrates the "goal" resolution (Figure 5.4).

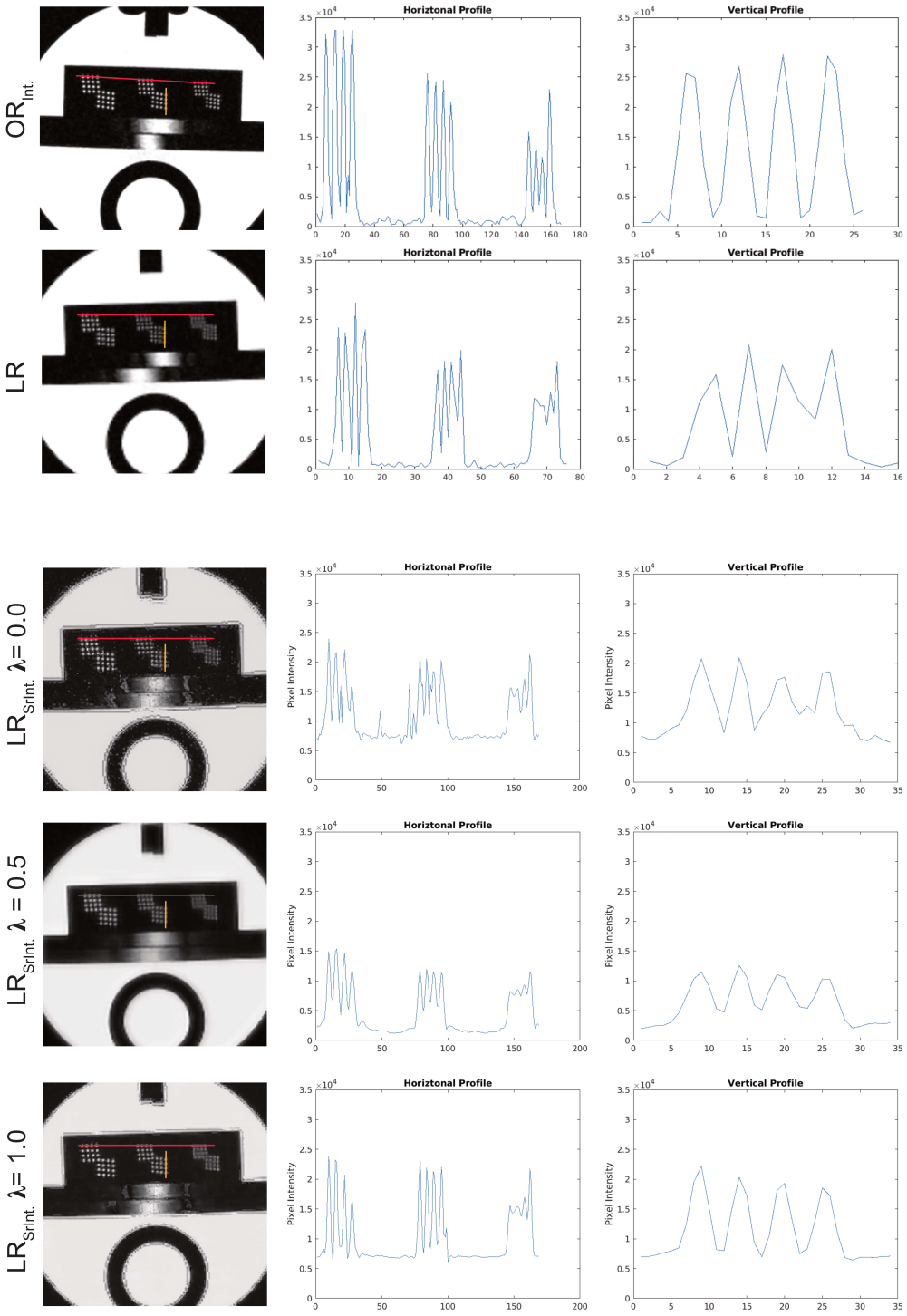


Figure 5.4: Images of the resolution insert in the small ACR phantom (first column). The second and third columns show the intensity profiles taken along the paths indicated by the red and yellow lines, respectively, in the corresponding image. The top two rows display results from the $OR_{Int.}$ and $LR$ acquisitions, while the bottom three rows show results from the SR pipeline using $W_{Interp.}$ with $\lambda = 0.0, 0.5,$ and 1.0.

Since the profiles resulting from the use of $\lambda = 0.5$ appeared to be the most similar to that of the $LR$ and $OR_{Int.}$ profiles, the same analysis was performed on $LR_{SrInt.}$ images resulting from the use of $W_{Interp.}$ with $\lambda = 0.4, 0.6,$ and 0.7. These results, along with those from $\lambda = 0.5$ are shown in Figure 5.5.

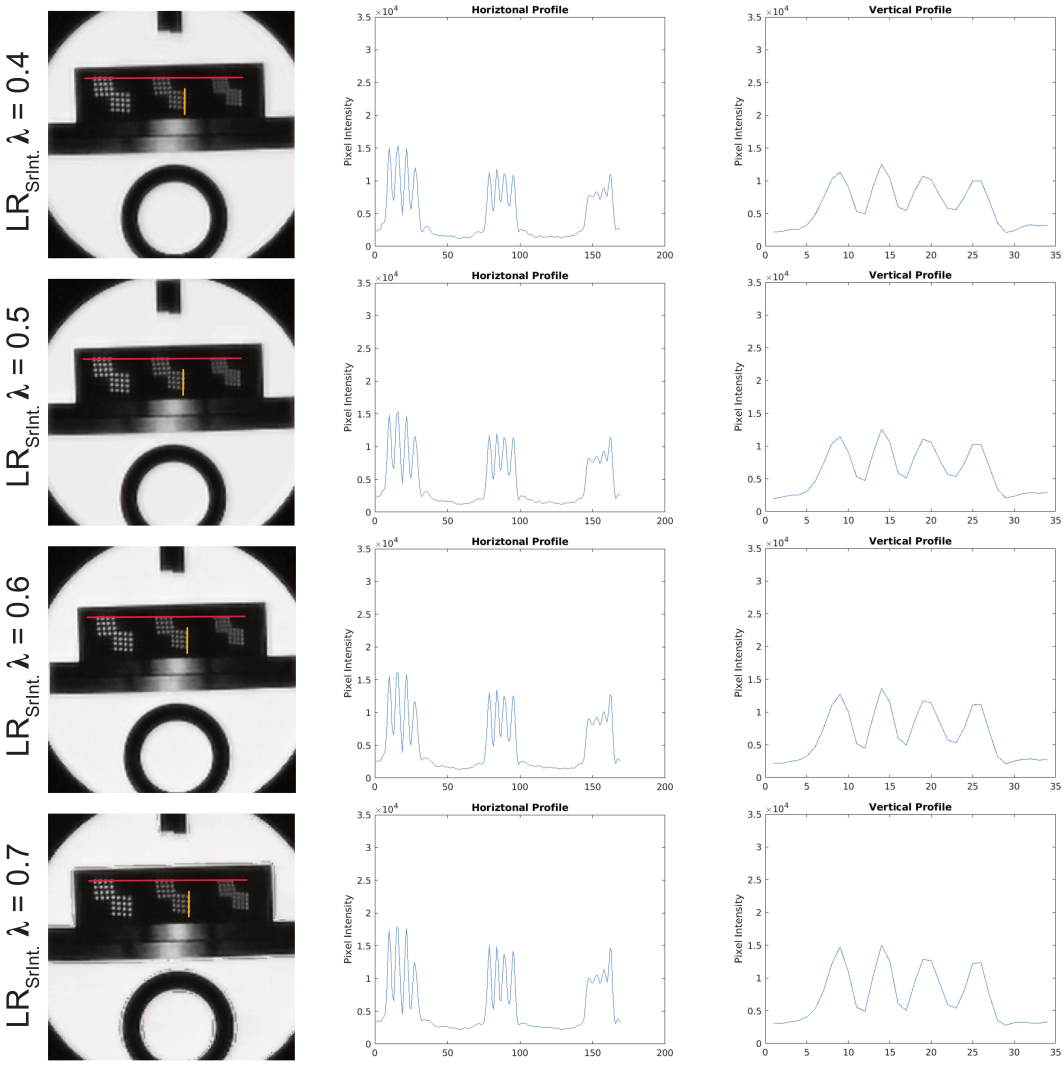


Figure 5.5: Images of the resolution insert in the small ACR phantom (first column). The second and third columns show the intensity profiles taken along the paths indicated by the red and yellow lines, respectively, in the corresponding image. Results are from the SR pipeline using $W_{Interp.}$ with $\lambda = 0.4, 0.5, 0.6,$ and 0.7.

For further evaluation, the SNR of each image was calculated via $SNR = 0.70\frac{S}{SD_{air}}$ where $S$ is the mean pixel intensity in an ROI drawn inside the $T_1$ contrast vial and $SD_{air}$ is the standard deviation of pixel intensities in an ROI drawn outside the

phantom in the lower left corner of the image. These regions were also used to calculate $CNR = 0.70\frac{S-S_{air}}{SD_{air}}$, where $S_{air}$ is the mean signal intensity in the air ROI, and fractional signal, $\frac{\Delta S}{S} = \frac{S-S_{air}}{S}$. A depiction of the regions used and the results of the calculations are shown in Figure 5.6.

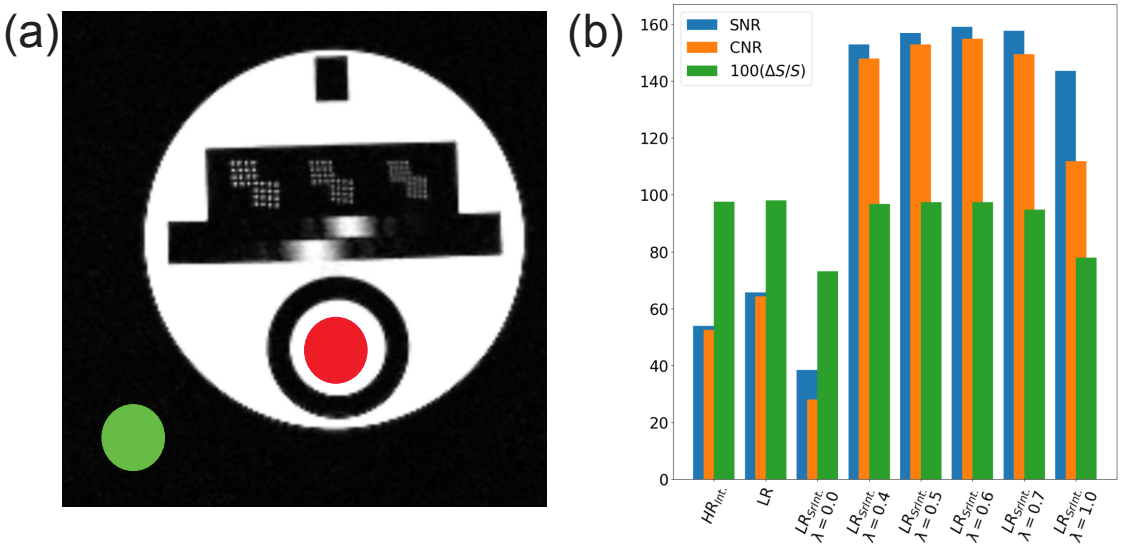


Figure 5.6: (a) Signal (red) and noise (green) ROIs used for calculation of (b) SNR, CNR, and $\Delta S/S$ in $OR_{Int.}$, and $LR$ acquisitions, and images resulting from the SR pipeline using $W_{Interp.}$ with $\lambda = 0.0, 0.4, 0.5, 0.6, 0.7,$ and 1.0

As is clear in Figures 5.4 and 5.6, the use of $W_{Interp.}$ with $\lambda = 0.0$ ($W_{PSNR}$) results in images that do not maintain the baseline contrast of the $OR_{Int.}$ and $LR$ acquisitions. $W_{NC}$ did improve the SNR and CNR of images compared to $OR_{Int}$ and $LR$, but the fractional signal is reduced. This indicates that the noise region has a lower standard deviation but a higher average signal, consistent with excessive "smoothing" required to reduce compression artifacts. Since the profiles in Figure 5.5 and the SNR, CNR, and $\Delta S/S$ in Figure 5.6 show comparable results for $\lambda = 0.4, 0.5, 0.6,$ and 0.7, we chose a value of $\lambda = 0.6$, to avoid the issues resulting from $\lambda = 0.0$ and 1.0.

A final summary of how the experimental data set for each participant ($OR_{Int.}$, $OR_{Sr}$, $LR_{Int.}$, and $LR_{SrInt.}$) was created is shown in Figure 5.7.

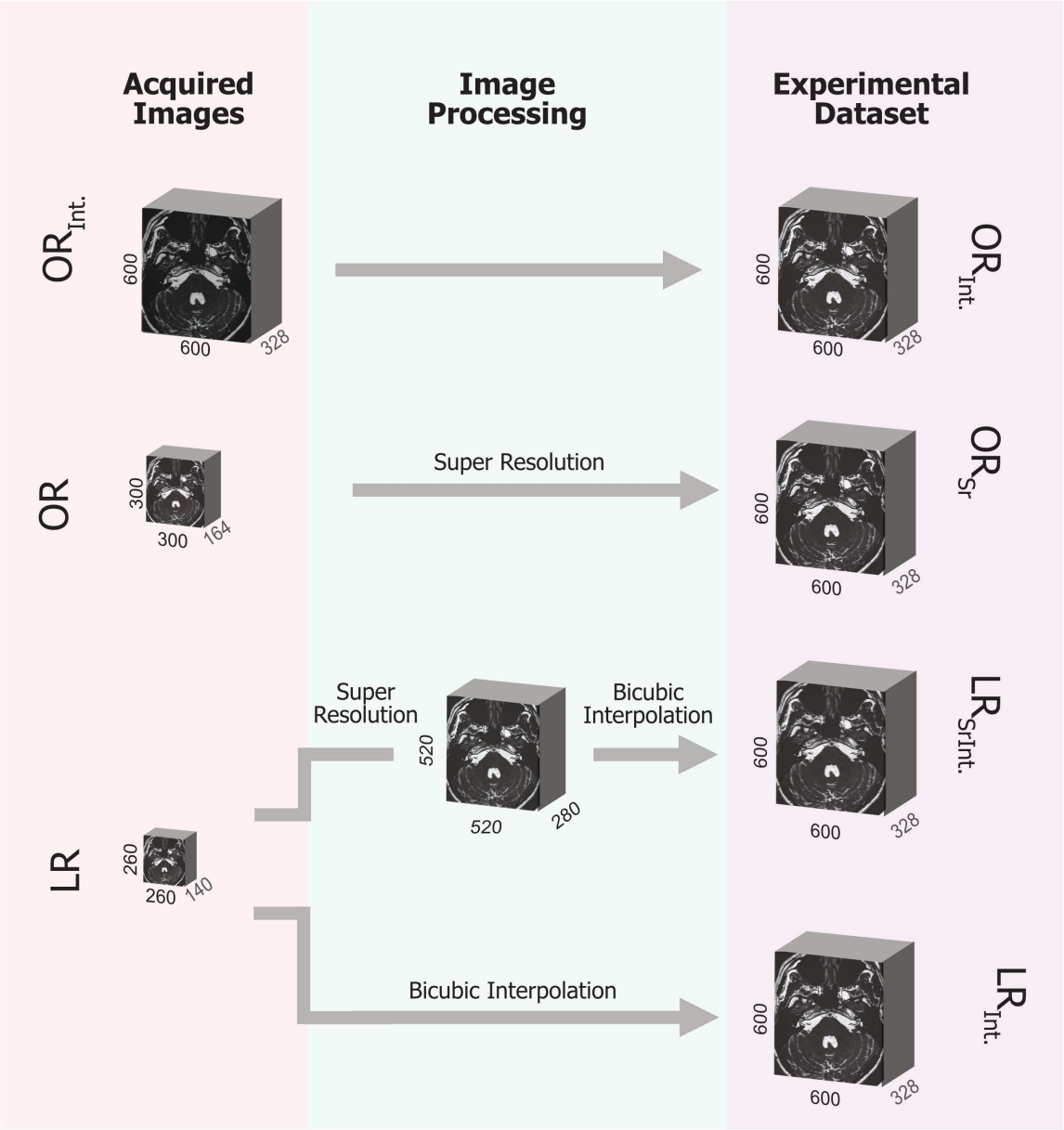


Figure 5.7: A summary of how the experimental data set for each participant was generated from the acquired images.

### 5.2.3 Image Analysis

The $OR_{Int.}$, $OR_{Sr}$, $LR_{Int.}$, and $LR_{SrInt.}$ images for each patient were converted to DICOM format with any identifying information removed, and shown to two-board certified radiologists (each with $>$ 5 years sub-speciality experience) in a randomized order. Raters were asked to independently rate their ability to visualize the superior

semi-circular canals (SSC), facial nerves, cochlear nerves, cochleas, and vestibules, as well as their impression of overall image quality. Ratings were done using a Likert scale ranging from 1 (very poor) to 5 (excellent). Ratings were recorded separately for left and right ears and ratings were subsequently concatenated such that there was one set of ratings per structure per rater. Gwet's AC2, with ordinal weighting, was calculated to determine inter-rater reliability for each structure under each image type.

Wilcoxon signed-rank tests were used to compare ratings between image types (as described in Section 5.1), for all structures examined as well as overall image quality. To correct for multiple comparisons of image data sets, a conservative Bonferroni correction was applied such that a $P$ value of $P < 0.02$ (i.e. $\approx \frac{0.05}{3}$) was considered significant. Python 3.6 and the Statistical Functions module were used to perform statistical calculations.

## 5.3 Results

A representative experimental data set from one participant is shown in Figure 5.8.

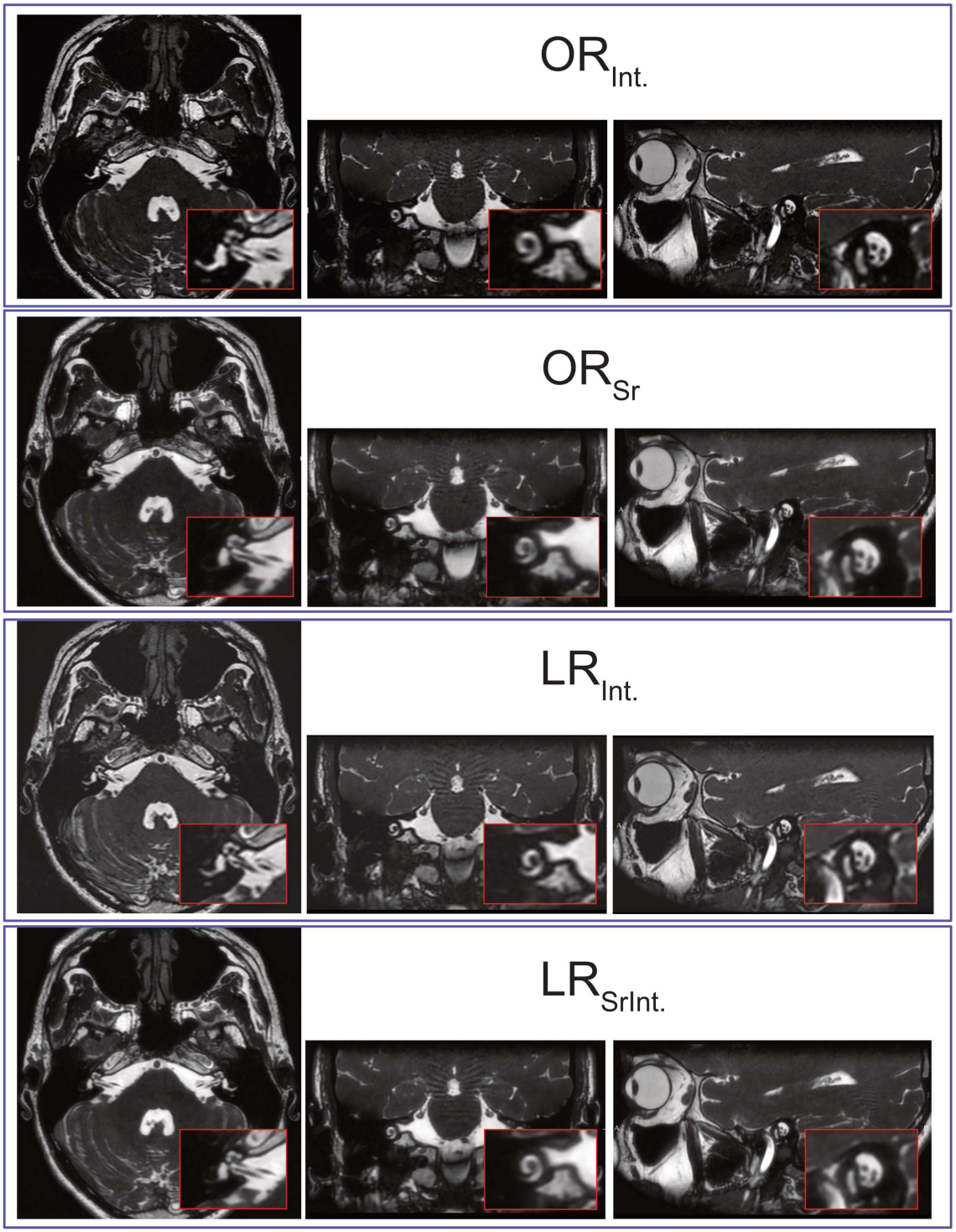


Figure 5.8: Example axial, coronal, and sagittal slices of an experimental data set from one participant. Coronal and sagittal slices have been cropped to remove excess background. Structures in the red boxes are zoomed-in versions of the vestibule, cochlea, and facial and cochlear nerves (axial), the cochlea (coronal), and cross sections of the facial and cochlear nerves (upper and lower left dark spots in the sagittal view).

The inter-rater reliability calculations (Table 5.3) yielded "poor" to "moderate" agreement between raters,[83] for all structures, for each of the image cohorts examined. As such, the ratings from each rater cannot be combined into a single data set, and instead will be considered as two unique case studies.

| | $OR_{Int.}$ | | $OR_{Sr}$ | | $LR_{Int.}$ | | $LR_{SrInt.}$ | |
|---|---|---|---|---|---|---|---|---|
| **Structure** | *AC2* | *95% C.I.* | *AC2* | *95% C.I.* | *AC2* | *95% C.I.* | *AC2* | *95% C.I.* |
| **SSC** | 0.01 | (-0.50, 0.52) | -0.42 | (-0.96, 0.12) | 0.19 | (-0.26, 0.65) | 0.25 | (-0.24, 0.74) |
| **Facial nerve** | -0.48 | (-0.95, -0.01) | -0.05 | (-0.62, 0.52) | 0.24 | (-0.20, 0.68) | -0.05 | (-0.51, 0.40) |
| **Cochlear nerve** | -0.10 | (-0.62, 0.41) | -0.30 | (-0.87, 0.27) | 0.28 | (-0.14, 0.70) | 0.24 | (-0.24, 0.72) |
| **Cochlea** | -0.14 | (-0.62, 0.34) | 0.21 | (-0.33, 0.76) | 0.16 | (-0.18, 0.51) | 0.18 | (-0.26, 0.63) |
| **Vestibule** | -0.25 | (-0.80, 0.30) | -0.12 | (-0.55, 0.31) | -0.21 | (-0.76, 0.33) | 0.29 | (-0.10, 0.68) |
| **Image quality** | -0.002 | (-0.78, 0.77) | 0.47 | (0.02, 0.92) | -0.38 | (-1.15, 0.39) | 0.14 | (-0.56, 0.84) |

Table 5.3: Gwet's AC2 values for inter-rater reliability and their corresponding 95% confidence intervals (95 % C.I.) for each structure examined in each type of image. Results show "poor" ($AC2 < 0$) to "fair" ($0.21 \leq AC2 \leq 0.40$) agreement between raters, apart from the image quality ratings of the $OR_{Sr}$ images which show "moderate" agreement ($0.41 \leq AC2 \leq 0.6$).

Results from the first rater are summarized in Figure 5.9, and those from the second rater are shown in Figure 5.10. For convenience, $P$ values resulting from all Wilxocon signed-rank tests are shown in Table 5.4.

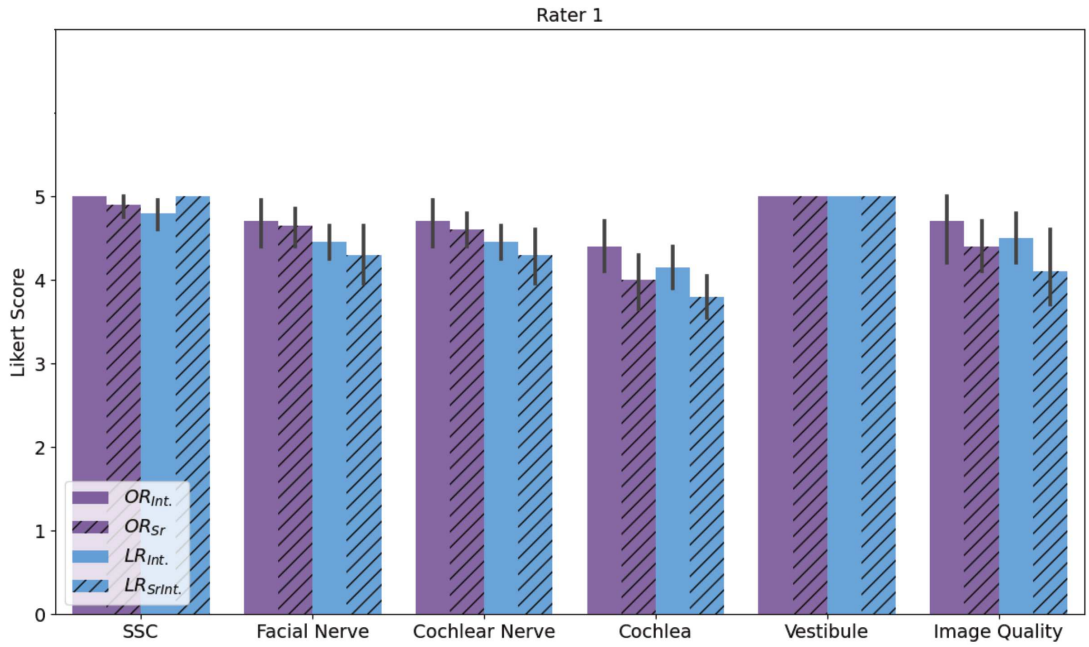


Figure 5.9: Average Likert scores from Rater 1. Error bars depict 95% confidence intervals.

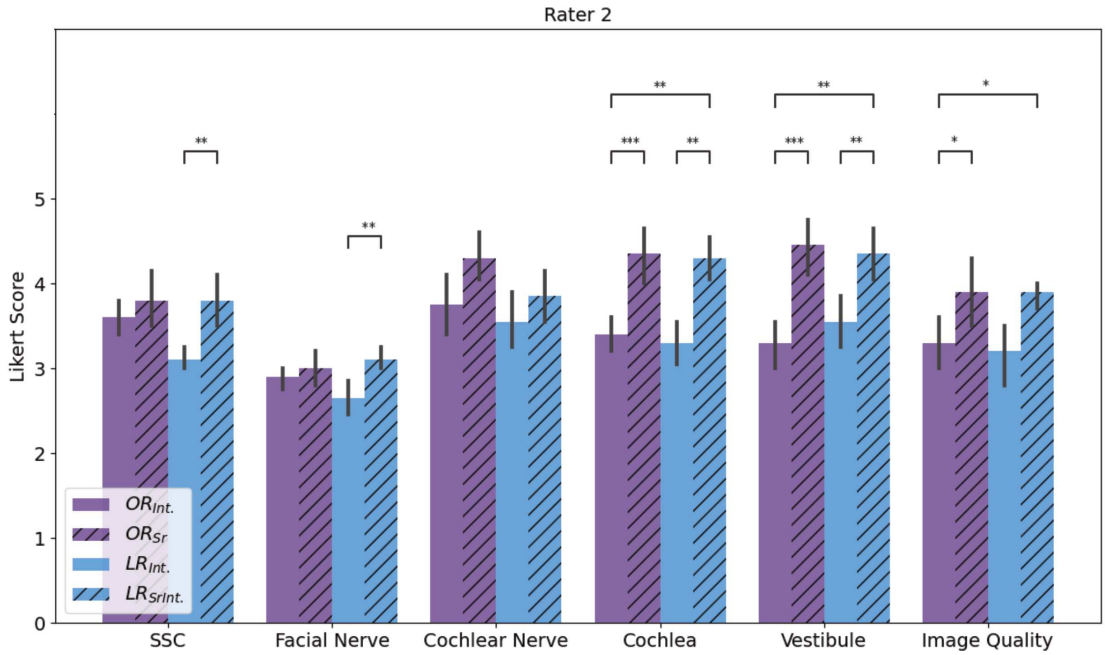


Figure 5.10: Average Likert scores for Rater 2. Error bars depict 95% confidence intervals. * indicates $0.003 < P \leq 0.02$, ** indicates $0.0003 < P \leq 0.003$, and *** indicates $0.00003 < P \leq 0.0003$.

| | $P$ Value From Wilcoxon Signed-Rank Test Rater 1 / Rater 2 | | |
|---|---|---|---|
| Structure | Comparison of $LR_{SrInt.}$ and $OR_{Int.}$ | Comparison of $LR_{Int.}$ and $LR_{SrInt.}$ | Comparison of $OR_{Int.}$ and $OR_{Sr}$ |
| SSC | 1.0 / 0.1 | 0.05 / **0.0005** | 0.16 /0.33 |
| Facial Nerve | 0.14 / 0.05 | 0.37 / **0.003** | 0.71 / 0.41 |
| Cochlear Nerve | 0.14 / 0.56 | 0.37 / 0.06 | 0.41 / 0.05 |
| Cochlea | 0.03 / **0.0006** | 0.12 / **0.001** | 0.05 / **0.0003** |
| Vestibule | 1.0 / **0.0004** | 1.0 / **0.002** | 1.0 / **0.0002** |
| Image Quality | 0.12 / **0.01** | 0.16 / 0.02 | 0.18 / **0.01** |

Table 5.4: $P$ values from Wilcoxon signed-rank tests. Values are presented in Rater 1 / Rater 2 format and bold values indicate a statistically significant difference between ratings, i.e. $P < 0.02$.

Analysis of ratings from the first rater yielded no statistically significant differences between ratings for any image type when visualizing any structures, nor for overall image quality.

Alternatively, there are a number of significant differences in the second rater's results. First, the $LR_{SrInt.}$ images were preferred over the $OR_{Int.}$ images for visualization of the cochlea, and the vestibule, as well as for overall image quality. There are no statistical differences when visualizing the SSC, facial nerve, or cochlear nerve. When comparing methods used to bring images to HR, the $LR_{SrInt.}$ images were preferred over the $LR_{Int.}$ images for visualizing the SSC, the facial nerve, the cochlea, and the vestibule, while no statistical difference was found for visualization of the cochlear nerve or for overall image quality. The $OR_{Sr}$ images are preferred over the $OR_{Int.}$ images for visualizing the cochlea, vestibule, and for overall image quality, and there are no statistical differences when visualizing the SSC, the facial nerve or the cochlear nerve.

## 5.4 Discussion

The objective of this study was to perform LR bSSFP acquisitions, allowing for the advantages of e.g. reduced scan times and reduced patient motion. Notably, the RBW used for the LR protocol was identical to that in the OR protocol so as to maintain comparable SNR in the resulting images, and a further reduction in TR

would be feasible if this were not a concern.

The fact that ISR was not designed, or trained, to operate on MR images was known prior to its use, and is therefore not considered a limitation of the study itself. It did require, however, that numerous pre- and post-processing steps were performed on the MR images, and thus the results are not necessarily representative of what the ISR model is capable of, nor of what is possible with SR of MRI volumes.

Despite this, however, neither radiologist's ratings of their ability to visualize structures of the temporal bone in HR images produced with the SR pipeline were significantly lower than HR images generated with standard interpolation methods. Notably, although the $LR$ images required acquisition times that were less than the $OR$ acquisitions, the time required to produce $LR_{SrInt.}$ images with the SR pipeline is substantially longer than that required for native reconstruction and interpolation by the system. Future investigations could focus on improving both acquisition and reconstruction times, potentially by using a model designed for 3D image volumes, thereby reducing the need for time-consuming pre- and post-processing steps.

The result that ratings from Raters 1 and 2 yielded poor to fair agreement is, itself, an important and compelling result. Subsequent discussions with Rater 1 revealed a preference for standard interpolation methods due to the $OR_{Sr}$ and $LR_{SrInt.}$ images appearing "over-smoothed". This is in contrast to Rater 2's results which, overall, show that the images generated with the SR pipeline were preferred over, or not rated significantly lower than, those generated with standard interpolation methods. While some MRI SR studies have shown similar levels of inter-rater reliability (e.g.[91]), others show much higher levels of agreement (e.g.[92]), yet little attention has been paid to this phenomenon in the literature. Given that MRI system vendors provide filtering options in their reconstruction pipelines to satisfy local preferences, we suspect that differences in training are responsible for the preferences demonstrated, and further investigation with more than two raters and a larger number of participants is required.

Overall, neither radiologist rated the $LR_{SrInt.}$ images significantly lower than $OR_{Int.}$ images when visualizing structures of the inner ear, nor for overall image quality, thereby permitting a reduction in scan time of 1 minute and 8 seconds.

### 5.4.1 Limitations

One main limitation of this study is the fact that we could never truly blind raters to image type, due to their experience, and bias may have therefore affected their ratings. Furthermore, no "calibration" session was performed for radiologists' use of the Likert scale, therefore experience may not only play an important role preference for image type, but in range of scores assigned to each image.

## 5.5 Conclusions

Since no significant differences were found between ratings provided by Rater 1, the first hypothesis, which predicted that LR images processed with the SR pipeline would not be rated as significantly lower than OR images processed with interpolation, can be confirmed for Rater 1. Hypotheses 2, that LR images processed with the SR pipeline would be rated significantly higher than those processed with interpolation, and hypothesis 3, that OR images processed with the SR pipeline would be rated significantly higher than those processed with interpolation, are both rejected for Rater 1.

The ratings from Rater 2 allow us to confirm hypothesis 1 for all structures examined as well as overall image quality. Hypothesis 2 is confirmed for the SSC, facial nerve, cochlea, and vestibule, and it is rejected for the cochlear nerve and for overall image quality. Finally, hypothesis 3 is confirmed for the cochlea, vestibule, and for overall image quality, and is rejected for the SSC, facial nerve, and cochlear nerve.

# Chapter 6

# Conclusions and Future Work

To our group's knowledge, the studies presented in this thesis are the first to provide valuable insight into the use of a 0.5T head-only MR system with high-strength, fast slew rate gradients, for effective and efficient bSSFP imaging of the challenging regions of the temporal bone and paranasal sinuses.

The first study (Chapter 3), showed that a bSSFP sequence could be used to image the temporal bone and paranasal sinuses and that the majority ($> 50\%$) of the resulting images, from both protocols, were free of artifacts traversing relevant structures. Unexpectedly, the non-phase cycled 0.5T temporal bone images provided superior artifact tolerance compared to 3T, phase-cycled acquisitions reconstructed with MIP, demonstrating a clear advantage to imaging regions with large susceptibility differences at low field strength. Both protocols achieved an average SNR greater than that preferred for visualization by radiologists,[76] at resolutions comparable to clinical acquisitions. In addition, the scan time of the 0.5T temporal bone acquisitions were over a minute shorter than the shortest clinical acquisition included in this study. The inability to compare SNR between the 0.5T and 3T temporal bone images presents a possible avenue for future work - describing the SNR of images acquired with MIP would allow for meaningful comparisons to be made. While the sinus protocol produced sufficient artifact tolerance, future adjustments to the protocol could further reduce the presence of banding artifacts by reducing resolution to allow for an even shorter TR.

The advantages of temporal bone bSSFP imaging at 0.5T were examined in a clinical context in the second study (Chapter 4). The improved artifact tolerance, established in the first study, led radiologists to prefer bSSFP temporal bone images acquired at 0.5T over conventional bSSFP acquisitions for visualization of the SSC and vestibule. While conventional acquisitions were preferred to visualize the facial

and cochlear nerves, no preference was found for visualizing the cochlea, nor when ratings from all structures were combined into an overall comparison. To further examine the clinical use of our 0.5T temporal bone protocol, studies that investigate sensitivity and specificity when using the images for diagnosis must be performed. Furthermore, these questions should be asked in the context of the entire temporal bone protocol, and not solely the bSSFP sequence. Temporal bone exams, including the clinical exams performed within our health institution, typically include a diffusion-weighted imaging sequence (DWI) to rule out the presence of vestibulocochlear schwannomas or cholesteatomas.[50] A DWI sequence should therefore be developed and included in the 0.5T temporal bone protocol for any further evaluations of clinical use. Although images resulting from the sinus protocol were not included in this study, there is a current study underway in our group in which the confidence of radiologists in answering clinical questions is assessed for both the 0.5T MRI protocol and current SOC CT acquisitions, and we expect the results to provide valuable feedback on future versions of the 0.5T bSSFP paranasal sinus protocol.

The exploration performed in the third study (Chapter 5) yielded a number of interesting results. Most importantly, we found that bringing low resolution 0.5T temporal bone bSSFP images to high resolution with a super resolution pipeline produced images that radiologists did not rate as significantly different from original resolution acquisitions brought to high resolution with standard interpolation methods, when visualizing structures of the inner ear. We can therefore state that the advantages conferred by low resolution bSSFP imaging were realized, without a significant degradation in image quality. More specifically, the super resolution pipeline permitted a reduction in scan time of 1 minute and 8 seconds compared to the original resolution acquisition. As discussed, reducing the resolution of the paranasal sinus bSSFP protocol would permit a shorter TR and improved artifact suppression. The results of the SR study were encouraging, and present an opportunity to obtain LR images of the paranasal sinuses, with improved artifact tolerance, that can be brought back to clinical resolutions without significant degradation. Finally, the pipeline presented in this work was limiting in that we were required to modify the input and output for use in our application. Additional future work could focus on training the ISR

model on MR images with the current pre- and post- processing steps applied, or alternatively, work could be done to develop our own model to perform the SR task on 3D MRI volumes.

The results of the studies presented herein demonstrate the new opportunities afforded by bSSFP imaging on modern, low-field MR systems. Leveraging the advantages of low-field imaging with high-strength gradients, and combining this with super resolution techniques, has the potential to make MR imaging of challenging regions not only possible, but fast and artifact-free, thereby improving patient experience without sacrificing image utility.

# Appendix A

## Derivation of the steady-state transverse magnetization

In a steady-state, the transverse magnetization, $\vec{M}$, does not change within one TR, i.e.

$$\frac{d\vec{M}}{dt} \cdot \vec{M} = 0 \tag{A.1}$$

where $\frac{d\vec{M}}{dt}$ is (i.e. Equation 2.4):

$$\begin{aligned}\frac{d\vec{M}}{dt} &= \vec{M} \times \gamma\vec{B} + \frac{(M_0 - M_z)\hat{z}}{T_1} - \frac{(M_x\hat{x} + M_y\hat{y})}{T_2} \\ &= \gamma M_y B_z \hat{x} - \gamma M_x B_z \hat{y} - \frac{M_x}{T_2}\hat{x} - \frac{M_y}{T_2}\hat{y} + \frac{M_0 - M_z}{T_1}\hat{z}\end{aligned}$$

for $\vec{B} = B_z\hat{z}$. Using $\vec{M} = M_x\hat{x} + M_y\hat{y} + M_z\hat{z}$, we find

$$\begin{aligned}&\frac{d\vec{M}}{dt} \cdot \vec{M} \\ &= \left[\left(\gamma M_y B_z - \frac{M_x}{T_2}\right)\hat{x} - \left(\gamma M_x B_z + \frac{M_y}{T_2}\right)\hat{y} + \left(\frac{M_0 - M_z}{T_1}\right)\hat{z}\right] \cdot (M_x\hat{x} + M_y\hat{y} + M_z\hat{z}) \\ &= \gamma M_x M_y B_z - \frac{M_x^2}{T_2} - \gamma M_x M_y B_z - \frac{M_y^2}{T_2} + \frac{M_0 - M_z}{T_1}M_z \\ &= \frac{-M_x^2 - M_y^2}{T_2} + \frac{M_0 - M_z}{T_1}M_z\end{aligned}$$

and from Equation A.1, we have

$$
\begin{aligned}
0 &= \frac{-M_x^2 - M_y^2}{T_2} + \frac{M_0 - M_z}{T_1} M_z \\
0 &= \frac{-T_1}{T_2}(M_x^2 + M_y^2) + M_0 M_z - M_z^2 \\
M_z^2 - M_0 M_z &= \frac{-T_1}{T_2}(M_x^2 + M_y^2) \\
\left(M_z - \frac{M_0}{2}\right)^2 - \left(\frac{M_0}{M_2}\right)^2 &= \frac{-T_1}{T_2}(M_x^2 + M_y^2) \\
\left(M_z - \frac{M_0}{2}\right)^2 + \frac{M_x^2 + M_y^2}{T_2/T_1} &= \left(\frac{M_0}{2}\right)^2 \\
\frac{\left(M_z - \frac{M_0}{2}\right)^2}{(\frac{M_0}{2})^2} + \frac{M_x^2 + M_y^2}{\frac{T_2}{T_1}(\frac{M_0}{2})^2} &= 1
\end{aligned}
$$

Now, considering the fact that $M_x^2 + M_y^2 = M_{xy}^2$, and letting $M_z = M\cos(\beta/2)$, and $M_{xy} = M\sin(\beta/2)$, we find

$$
\begin{aligned}
\left[M\cos\left(\frac{\beta}{2}\right) - \frac{M_0}{2}\right]^2 + \frac{M^2\sin^2(\frac{\beta}{2})}{\frac{T_2}{T_1}} &= \left(\frac{M_0}{2}\right)^2 \\
M^2\cos^2\left(\frac{\beta}{2}\right) - M_0 M\cos\left(\frac{\beta}{2}\right) + \left(\frac{M_0}{2}\right)^2 + \frac{T_1}{T_2}M^2\sin^2\left(\frac{\beta}{2}\right) &= \left(\frac{M_0}{2}\right)^2 \\
M^2\cos^2\left(\frac{\beta}{2}\right) + \frac{T_1}{T_2}M^2\sin^2\left(\frac{\beta}{2}\right) &= M_0 M\cos\left(\frac{\beta}{2}\right) \\
M\cos^2\left(\frac{\beta}{2}\right) + \frac{T_1}{T_2}M\sin^2\left(\frac{\beta}{2}\right) &= M_0\cos\left(\frac{\beta}{2}\right)
\end{aligned}
$$

multiplying through by $\sin\left(\frac{\beta}{2}\right)$:

$$
\begin{aligned}
\sin\left(\frac{\beta}{2}\right)\left[M\cos^2\left(\frac{\beta}{2}\right) + \frac{T_1}{T_2}M\sin^2\left(\frac{\beta}{2}\right)\right] &= \sin\left(\frac{\beta}{2}\right) M_0\cos\left(\frac{\beta}{2}\right) \\
M\sin\left(\frac{\beta}{2}\right) &= \frac{M_0\sin\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right)}{\cos^2\left(\frac{\beta}{2}\right) + \frac{T_1}{T_2}\sin^2\left(\frac{\beta}{2}\right)}
\end{aligned}
$$

Finally, dividing both the numerator and denominator by $\sin\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right)$,

$$M\sin\left(\frac{\beta}{2}\right) = \frac{M_0}{\frac{\cos^2\left(\frac{\beta}{2}\right)}{\sin\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right)} + \frac{T_1}{T_2}\frac{\sin^2\left(\frac{\beta}{2}\right)}{\sin\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right)}}$$

$$M_{xy} = \frac{M_0}{\cot\left(\frac{\beta}{2}\right) + \frac{T_1}{T_2}\tan\left(\frac{\beta}{2}\right)}$$

# Bibliography


[1] Edwin J.R. van Beek, Christiane Kuhl, Yoshimi Anzai, Patricia Desmond, Richard L. Ehman, Qiyong Gong, Garry Gold, Vikas Gulani, Margaret Hall-Craggs, Tim Leiner, C. C.Tschoyoson Lim, James G. Pipe, Scott Reeder, Caroline Reinhold, Marion Smits, Daniel K. Sodickson, Clare Tempany, H. Alberto Vargas, and Meiyun Wang. Value of mri in medicine: More than just another test? *Journal of Magnetic Resonance Imaging*, 49:e14–e25, 6 2019.

[2] Val M. Runge and Johannes T. Heverhagen. Advocating the development of next-generation, advanced-design low-field magnetic resonance systems. *Investigative Radiology*, 55:747–753, 12 2020.

[3] Hans-Martin Klein. Low-field magnetic resonance imaging. *Fortschr Röntgenstr*, 192:537–548, 2020.

[4] Mathieu Sarracanie and Najat Salameh. Low-field mri: How low can we go? a fresh view on an old debate. *Frontiers in Physics*, 8, 6 2020.

[5] Giuseppe BarisAno, Farshid Sepehrband, Kay Jann, Ryan Cabeen, Danny J Wang, Arthur Toga, and Meng Law. clinical 7 t mri: Are we there yet? a review about magnetic resonance imaging at ultra-high field, 2019.

[6] Mathias Blasche. Gradient performance and gradient amplifier power. *MAGNETOM Flash*, 69, 2017.

[7] Jeff A Stainsby, Geron A Bindseil, Ian Ro Connell, Gilbert Thevathasan, Andrew T Curtis, Phil J Beatty, Chad T Harris, Curtis N Wiens, and Alex Panther. Imaging at 0.5 t with high-performance system components system components. ISMRM, 5 2020.

[8] Franz Schmitt, Stefan Nowak, and Eva Eberlein. An attempt to reconstruct the history of gradient-system technology at siemens. *MAGNETOM Flash*, 77, 2020.

[9] Daniel Hernandez and Kyoung-Nam Kim. A review on the rf coil designs and trends for ultra high field magnetic resonance imaging. *Investigative Magnetic Resonance Imaging*, 24:95, 2020.

[10] David Lurie, Silvia Casini, and Neil Curtis. Mark-1: The world's first whole body scanner, 6 2016. Grampian Hospitals Art Trust.

[11] Bernhard Gruber, Martijn Froeling, Tim Leiner, and Dennis W.J. Klomp. Rf coils: A practical guide for nonphysicists. *Journal of Magnetic Resonance Imaging*, 48:590–604, 9 2018.

[12] David A. Feinberg, James D. Hale, Jeffrey C. Watts, Leon Kaufman, and Alexander Mark. Halving mr imaging time by conjugation: Demonstration at 3.5kg. *Radiology*, 161:527–531, 1986.

[13] Mark A. Griswold, Peter M. Jakob, Robin M. Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase. Generalized autocalibrating partially parallel acquisitions (grappa). *Magnetic Resonance in Medicine*, 47:1202–1210, 2002.

[14] Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, and Peter Boesiger. Sense: Sensitivity encoding for fast mri. *Magn Reson Med*, 42:952–962, 1999.

[15] Robin M. Heidemann, Özkan Özsarlak, Paul M. Parizel, Johan Michiels, Berthold Kiefer, Vladimir Jellus, Mathias Müller, Felix Breuer, Martin Blaimer, Mark A. Griswold, and Peter M. Jakob. A brief review of parallel magnetic resonance imaging. *European Radiology*, 13:2323–2337, 10 2003.

[16] Michael Lustig, David Donoho, and John M. Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine*, 58:1182–1195, 12 2007.

[17] Mehmet Akçakaya, Steen Moeller, Sebastian Weingärtner, and Kâmil Uğurbil. Scan-specific robust artificial-neural-networks for k-space interpolation (raki) reconstruction: Database-free deep learning for fast imaging. *Magnetic Resonance in Medicine*, 81:439–453, 1 2019.

[18] Shanshan Wang, Zhenghang Su, Leslie Ying, Xi Peng, Shun Zhu, Feng Liang, Dagan Feng, and Dong Liang. Accelerating magnetic resonance imaging via deep learning. pages 514–517. IEEE Computer Society, 6 2016.

[19] Chang Min Hyun, Hwa Pyung Kim, Sung Min Lee, Sungchul Lee, and Jin Keun Seo. Deep learning for undersampled mri reconstruction. *Physics in Medicine and Biology*, 63, 6 2018.

[20] Bo Zhu, Jeremiah Z. Liu, Stephen F. Cauley, Bruce R. Rosen, and Matthew S. Rosen. Image reconstruction by domain-transform manifold learning. *Nature*, 555:487–492, 3 2018.

[21] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. http://arxiv.org/abs/1501.00092, 12 2014.

[22] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. http://arxiv.org/abs/1512.03385, 12 2015.

[23] Tong Tong, Gen Li, Xiejie Liu, and Qinquan Gao. Image super-resolution using dense skip connections. volume 2017-October, pages 4809–4817. Institute of Electrical and Electronics Engineers Inc., 12 2017.

[24] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network. volume 2017-January, pages 105–114. Institute of Electrical and Electronics Engineers Inc., 11 2017.

[25] Chi Hieu Pham, Aurelien Ducournau, Ronan Fablet, and Francois Rousseau. Brain mri super-resolution using deep 3d convolutional networks. pages 197–200. IEEE Computer Society, 6 2017.

[26] Yuhua Chen, Yibin Xie, Zhengwei Zhou, Feng Shi, Anthony G. Christodoulou, and Debiao Li. Brain mri super resolution using 3d deep densely connected neural networks. volume 2018-April, pages 739–742. IEEE Computer Society, 5 2018.

[27] Yuhua Chen, Feng Shi, Anthony G. Christodoulou, Yibin Xie, Zhengwei Zhou, and Debiao Li. Efficient and accurate mri super-resolution using a generative adversarial network and 3d multi-level densely connected network. volume 11070 LNCS, pages 91–99. Springer Verlag, 2018.

[28] Nicolas Behl. Deep resolve-mobilizing the power of networks. *MAGNETOM Flash*, 78, 2021.

[29] Andre Fischer. Revisiting the physics behind mri and the opportunities that lower field strengths offer. *Magnetom FLASH*, 78, 1 2021.

[30] José P. Marques, Frank F.J. Simonis, and Andrew G. Webb. Low-field mri: An mr physics perspective. *Journal of Magnetic Resonance Imaging*, 49:1528–1542, 6 2019.

[31] Alex Panther, Gilbert Thevathasen, Ian R O Connell, Yi Yao, Curtis N Wiens, Andrew T Curtis, Geron A Bindseil, Chad T Harris, Philip J Beatty, Jeff A Stainsby, Charles H Cunningham, B A Chronik, and Cameron Piron. A dedicated head-only mri scanner for point-of-care imaging barriers to mri being used at the point-of-care. ISMRM, 5 2020.

[32] Stephan Biber. How to make it big inside and small outside. *Magnetom FLASH*, 78, 1 2020.

[33] Jana G. Delfino, Daniel M. Krainak, Stephanie A. Flesher, and Donald L. Miller. Mri-related fda adverse event reports: A 10-yr review. *Medical Physics*, 46:5562–5571, 12 2019.

[34] Adrienne E. Campbell-Washburn, Rajiv Ramasawmy, Matthew C. Restivo, Ipshita Bhattacharya, Burcu Basar, Daniel A. Herzka, Michael S. Hansen, Toby Rogers, W. Patricia Bandettini, Delaney R. McGuirt, Christine Mancini, David Grodzki, Rainer Schneider, Waqas Majeed, Himanshu Bhat, Hui Xue, Joel Moss,

Ashkan A. Malayeri, Elizabeth C. Jones, Alan P. Koretsky, Peter Kellman, Marcus Y. Chen, Robert J. Lederman, and Robert S. Balaban. Opportunities in interventional and diagnostic imaging by using high-performance low-field-strength mri. *Radiology*, 293:384–393, 2019.

[35] Seung Kyun Lee, Jean Baptiste Mathieu, Dominic Graziani, Joseph Piel, Eric Budesheim, Eric Fiveland, Christopher J. Hardy, Ek Tsoon Tan, Bruce Amm, Thomas K.F. Foo, Matt A. Bernstein, John Huston, Yunhong Shu, and John F. Schenck. Peripheral nerve stimulation characteristics of an asymmetric head-only gradient coil compatible with a high-channel-count receiver array. *Magnetic Resonance in Medicine*, 76:1939–1950, 12 2016.

[36] Chad T Harris, Andrew T Curtis, Curtis N Wiens, Philip J Beatty, and Jeff A Stainsby. Acoustic behavior of high performance imaging on a head-only 0.5t scanner with asymmetric gradients. ISMRM, 6 2019.

[37] Dmitriy A. Yablonskiy, Alexander L. Sukstanskii, and Joseph J.H. Ackerman. Image artifacts in very low magnetic field mri: The role of concomitant gradients. *Journal of Magnetic Resonance*, 174:279–286, 6 2005.

[38] Christopher T. Sica and Craig H. Meyer. Concomitant gradient field effects in balanced steady-state free precession. *Magnetic Resonance in Medicine*, 57:721–730, 2007.

[39] Yicun Wang, Peter van Gelderen, Jacco A. de Zwart, and Jeff H. Duyn. B0-field dependence of mri t1 relaxation in human brain. *NeuroImage*, 213, 6 2020.

[40] A Oppelt, R Graumann, H Barfuss, H Fischer, W Hartl, and W Schajor. Fisp - a new fast mri sequence. *Electromedia*, 54:15–18, 1986.

[41] O Bieri and G Bauman. S2.o9. high-resolution imaging at low-fields: how far can we go? *Magnetic Resonance Materials in Physics, Biology, and Medicine*, 34 (Suppl 1):S16–S17, 2021.

[42] Klaus Scheffler and Stefan Lehnhardt. Principles and applications of balanced ssfp techniques. *European Radiology*, 13:2409–2418, 11 2003.

[43] Govind B. Chavhan, Paul S. Babyn, Bhavin G. Jankharia, Hai Ling M. Cheng, and Manohar M. Shroff. Steady-state mr imaging sequences: Physics, classification, and clinical applications. *Radiographics*, 28:1147–1160, 3 2008.

[44] Karla L. Miller. Asymmetries of the balanced ssfp profile. part i: Theory and observation. *Magnetic Resonance in Medicine*, 63:385–395, 2010.

[45] Eric G. Ekdale. Form and function of the mammalian inner ear. *Journal of Anatomy*, 228:324–337, 2 2016.

[46] Joseph E. Hawkins. Human ear. https://www.britannica.com/science/ear, 9 2022.

[47] Amy F. Juliano, Daniel T. Ginat, and Gul Moonis. Imaging review of the temporal bone: Part i. anatomy and inflammatory and neoplastic processes. *Radiology*, 269:17–33, 10 2013.

[48] Samuel A Spear, M Jackson, Yzrahul Mehta, E Morel, S Miller, Zdwayne Anderson, and A Arriaga. Is mri equal to ct in the evaluation of thin and dehiscent superior semicircular canals? *Otol Neurotol*, 37:167–170, 2016.

[49] P. Browaeys, T. L. Larson, M. L. Wong, and U. Patel. Can mri replace ct in evalu. *American Journal of Neuroradiology*, 34:1421–1427, 7 2013.

[50] John C. Benson, Matthew L. Carlson, and John I. Lane. Mri of the internal auditory canal, labyrinth, and middle ear: How we do it. *Radiology*, 297:252–265, 11 2020.

[51] Na Young Jung, Won-Jin Moon, Min Hee Lee, and Eun Chul Chung. Magnetic resonance cisternography: Comparison between 3-dimensional driven equilibrium with sensitivity encoding and 3-dimensional balanced fast-field echo sequences with sensitivity encoding. *J Comput Assist Tomogr* , 31:588–591, 2007.

[52] National Cancer Institute. Paranasal sinuses. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/paranasal-sinus.

[53] Kolawole Okiyemi and Terance Tsue. Radiologic imaging in the management of sinusitis. *American Family Physician*, 66(10):1882–1886, 11 2002.

[54] Laura B. Eisenmenger and Yoshimi Anzai. *Acute Sinusitis in Adults and Children: Evidence-Based Emergency Imaging*, pages 183–203. Springer International Publishing, Cham, 2018.

[55] Barton F. Branstetter IV and Jane L. Weissman. Role of mr and ct in the paranasal sinuses. *Otolaryngologic Clinics of North America*, 38:1279–1299, 12 2005.

[56] Daniel H. Coelho, Taylor S. Pence, Mostafa Abdel-Hamid, and Richard M. Costanzo. Cribriform plate width is highly variable within and between subjects. *Auris Nasus Larynx*, 45:1000–1005, 10 2018.

[57] Thomas A. Gallagher, Alexander J. Nemeth, and Lotfi Hacein-Bey. An introduction to the fourier transform: Relationship to mri. *American Journal of Roentgenology*, 190:1396–1405, 5 2008.

[58] Brian Hargreaves. Rapid gradient-echo imaging. *Journal of Magnetic Resonance Imaging*, 36:1300–1313, 12 2012.

[59] Oliver Bieri and Klaus Scheffler. Fundamentals of balanced steady state free precession mri. *Journal of Magnetic Resonance Imaging*, 38:2–11, 7 2013.

[60] Andrew M. Elliott, Matt A. Bernstein, Heidi A. Ward, John Lane, and Robert J. Witte. Nonlinear averaging reconstruction method for phase-cycle ssfp. *Magnetic Resonance Imaging*, 25:359–364, 4 2007.

[61] Andrew D. Brown and Thomas R. Marotta. Using machine learning for sequence-level automated mri protocol selection in neuroradiology. *Journal of the American Medical Informatics Association*, 25:568–571, 5 2018.

[62] Farhad Malik. Neural networks bias and weights: Understanding the two most important concepts. https://medium.com/fintechexplained/neural-networks-bias-and-weights-10b53e6285da, 5 2019.

[63] Alexander Selvikvåg Lundervold and Arvid Lundervold. An overview of deep learning in medical imaging focusing on mri. *Zeitschrift fur Medizinische Physik*, 29:102–127, 5 2019.

[64] IBM Cloud Education. Convolutional neural networks. https://www.ibm.com/cloud/learn/convolutional-neural-networks, 10 2020.

[65] Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter. Fast and accurate deep network learning by exponential linear units (elus). 11 2015.

[66] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. http://arxiv.org/abs/1502.03167, 2 2015.

[67] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image super-resolution. http://arxiv.org/abs/1802.08797, 2 2018.

[68] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. http://arxiv.org/abs/1406.2661, 6 2014.

[69] Jason Brownlee. A gentle introduction to generative adversarial networks (gans). https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/, 6 2017.

[70] Kaiyu Yang, Klint Qinami, Li Fei-Fei, Jia Deng, and Olga Russakovsky. Towards fairer datasets: Filtering and balancing the distribution of the people subtree in the imagenet hierarchy. In *Conference on Fairness, Accountability, and Transparency*, 2020.

[71] Mariana Bento, Irene Fantini, Justin Park, Leticia Rittner, and Richard Frayne. Deep learning in large and multi-site structural brain mr imaging datasets. *Frontiers in Neuroinformatics*, 15, 1 2022.

[72] Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. Loss functions for image restoration with neural networks. https://github.com/NVlabs/PL4NN, 2017.

[73] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. http://arxiv.org/abs/1409.1556, 9 2014.

[74] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. http://arxiv.org/abs/1603.08155, 3 2016.

[75] Field R Howard Hart and Paul A Bottomley William A Edelstein Steven G Karr William M Leue Otward Mueller Rowland W Redington John F Schenck L Scott Smith Dimitrios Vatis. Nuclear magnetic resonance imaging: Contrast-to-noise ratio as a function of strength of magnetic. *Am J Roentgenol*, 141:1195–1201, 1983.

[76] Rodney S Owen and Felix W Wehrli. Original contribution predictability of snr and reader preference in clinical mr imaging. *Magnetic Resonance Imaging*, 8:737–745, 1990.

[77] Nema standards publication national electrical manufacturers association american national standard for characterization of phased array coils for diagnostic magnetic resonance images. www.nema.org, 2008.

[78] Allister Mason, James Rioux, Sharon E. Clarke, Andreu Costa, Matthias Schmidt, Valerie Keough, Thien Huynh, and Steven Beyea. Comparison of objective image quality metrics to expert radiologists' scoring of diagnostic quality of mr images. *IEEE Transactions on Medical Imaging*, 39:1064–1072, 4 2020.

[79] Kilem L. Gwet. Agreestat360. http://agreestat360.com, 2019.

[80] DV Cicchetti and AR Feinstein. High agreement but low kappa: Ii. resolving the paradoxes. *J Clin Epidemiol*, 43:551–558, 1990.

[81] MT Cibulka and MJ Strube. The conundrum of kappa and why some musculoskeletal tests appear unreliable despite high agreement: A comparison of cohen kappa and gwet ac to assess observer agreement when using nominal and ordinal data. *Phys Ther*, 101:1–5, 2021.

[82] N Wongpakaran, T Wongpakaran, D Wedding, and KL Gwet. A comparison of cohen's kappa and gwet's ac1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol*, 13, 2013.

[83] J Richard Landis and Gary G Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 3 1977.

[84] M. L. de Leeuw den Bouter, G. Ippolito, T. P. A. O'Reilly, R. F. Remis, M. B. van Gijzen, and A. G. Webb. Deep learning-based single image super-resolution for low-field mr brain images. *Scientific Reports*, 12:6362, 12 2022.

[85] Wang Z., Bovik A.C., Sheikh H.R., and Simoncelli E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.*, 13:600–612, 2007.

[86] Francesco Cardinale. Image super resolution. https://github.com/idealo/image-super-resolution, 2018.

[87] Wenzhe Shi, Jose Caballero, Ferenc Huszar, Johannes Totz, Andrew P. Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. volume 2016-December, pages 1874–1883. IEEE Computer Society, 12 2016.

[88] Andrey Ignatov, Radu Timofte, et al. Pirm challenge on perceptual image enhancement on smartphones: report. In *European Conference on Computer Vision (ECCV) Workshops*, January 2019.

[89] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, and Xiaoou Tang. Esrgan: Enhanced super-resolution generative adversarial networks. http://arxiv.org/abs/1809.00219, 9 2018.

[90] American College of Radiology. Phantom test guidance for use of the small mri phantom for the mri accreditation program. Small Phantom Guidance, 4 2018.

[91] Jeffrey D. Rudie, Tyler Gleason, Matthew J. Barkovich, David M. Wilson, Ajit Shankaranarayanan, Tao Zhang, Long Wang, Enhao Gong, Greg Zaharchuk, and Javier E. Villanueva-Meyer. Clinical assessment of deep learning–based super-resolution for 3d volumetric brain mri. *Radiology: Artificial Intelligence*, 4, 3 2022.

[92] Akshay S. Chaudhari, Zhongnan Fang, Feliks Kogan, Jeff Wood, Kathryn J. Stevens, Eric K. Gibbons, Jin Hyung Lee, Garry E. Gold, and Brian A. Hargreaves. Super-resolution musculoskeletal mri using deep learning. *Magnetic Resonance in Medicine*, 80:2139–2154, 11 2018.