NIRAL Guide to

MRI Quality Control Rating

Purpose	The document is intended to define the method for determining and tracking the quality of MRI scans acquired in research studies. It employs a case-by-case rating scale, Pass/Fail marker and artifact tracking system and describes how to apply these tools to structural Magnetic Resonance Images (sMRI). This manual focuses on T1 and T2 images and is intended to enable consistent, reproducible QC results which can be effectively communicated. This rating process is intended for use with grey level images.
Scope	This document is intended for use by Research Technicians within the Neuro-Image Analysis Lab and associated labs at the University of North Carolina at Chapel Hill. Prior knowledge of MR Images, brain anatomy and computers is assumed.

I. Introduction

The NIRAL lab rates MRI images as part of a quality control system that ensures the integrity of MRI data. Experienced research technicians gauge the usefulness of each MRI image for further processing, excluding those images that would yield unreliable data. Raters base these exclusions on the presence and severity of artifacts within the image.
This document includes:

- An overview of the rating scale and artifact tracking system
- A guide to the use of the system, with examples
- Steps for recording ratings and an example spreadsheet
- Common terms that are useful for discussions around the rating system

Note: Visualization Software - Raters within the NeuroImage Analysis and Research Laboratory use InsightSnap for visualization of images for QC. InsightSnap is freely available at the following website:

http://www.itksnap.org/pmwiki/pmwiki.php?n=Main.Downloads

II. Overview of the Quality Scale: 6-point Scale

This QC system employs two quality rating categories and an artifact tracking system which determines the suitability of MRI images for further processing. A narrative 6-point scale (or 4-point scale) provides an overall image quality rating, while an easily searchable binary code (0 - Fail or 1 – Pass; or P/F) indicates a basic Pass/Fail decision for each image. Additionally, each image is rated for types of defects or artifacts: motion, eye motion, hemodynamic and coverage. Examples of these artifacts are presented in section V of this document. Artifacts are tracked on the QC spreadsheet giving detailed information to researchers about the types of defects each image may possess.

6-point Scale

Technicians employ the 6-point scale to rate the quality of each image. In general, the rating scale can be thought of as a continuum from 1-6. Increasing numbers correspond to increasing severity or persistence of artifacts. The highest possible rating is "1" for "good". The lowest possible rating is "6" for "bad". The "1" or "good" rating is reserved for those scans that are of truly exceptional quality while the average scan is typically rated at "2" or "3". A scan rated "6" is considered unusable for further processing.
Several factors combine to determine the quality of each image, including:

- severity of artifacts
- estimated number of slices in the image which contain artifacts
- rough distribution of artifacts throughout the image

The following table outlines how each rating aligns with artifact density and location/distribution to create broad, yet functional markers.

Rating	Meaning	Artifacting	Location/Distribution
1	Good	No artifacts detected	throughout the image
2	Minor	Slight artifacting	restricted to a slice or two in the image
3	Usable	Moderate artifacting	on a slice or two
		Minor artifacting	on many slices throughout the image
4	Conditional	Heavy artifacting	on a slice or two
		Moderate artifacting	on many slices throughout the image
5	Poor	Heavy artifacting	on more than two slices but not throughout
		Moderate artifacting	throughout the image
6	Bad	Heavy artifacting	throughout the image

Note: When examining study scan data, it is important to remember that image quality control is inexact in its nature. The simple presence of artifacts does not necessarily exclude an image.

Pass/Fail

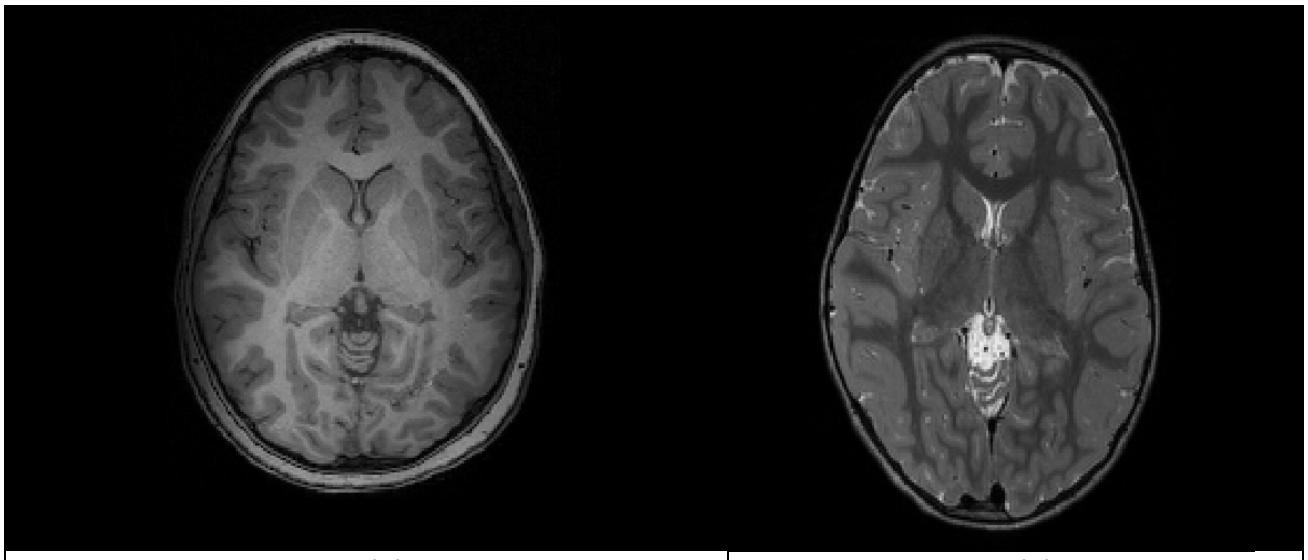
Each image is given a 1 or a 0 to indicate whether the scan quality is sufficient. Images rated as "Poor" or "Bad" do not produce reliable data even if a processing result is obtained. Therefore, raters generally exclude these images from further processing by incorporating a "0" (Fail) as a usability measure. The extra notation is easily searchable and can provide an override for borderline cases.

Artifact Tracking

Finally, technicians identify 4 types of artifacts in each image; motion, eye motion, hemodynamic and coverage. The presence or absence of each type of artifact is recorded in specific columns for each image, If the image possesses a given artifact, a 1 is assigned to that column. If the image does not possess a given artifact, a 0 is assigned to that column. The artifacts are discussed in detail in Section V of this document.

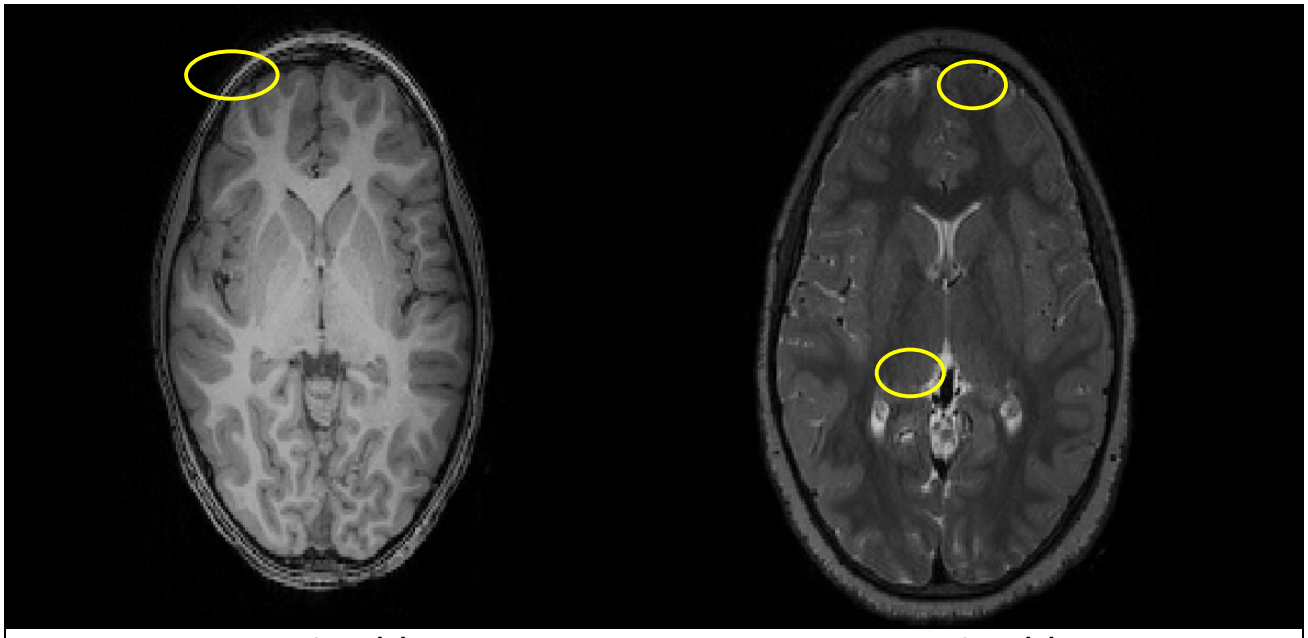
III. Image Category Examples: Using a 6-point Scale

The following screenshots illustrate the 6-point rating scale for both T1 and T2 images. While these images are viewed and evaluated in all three dimensions, the screenshots are shown in the axial plane for simplicity.



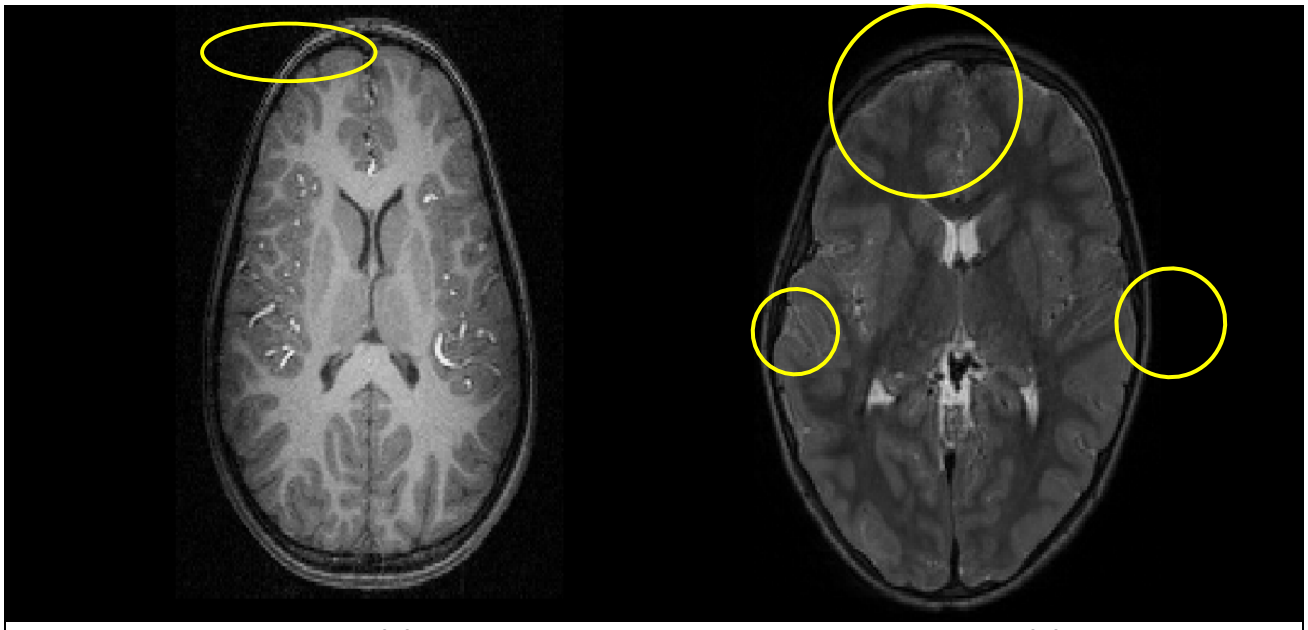
Good (1) T1	Good (2) T2

1 – Good category. No artifacts detected throughout image. These images are free from any visually detectable defect.



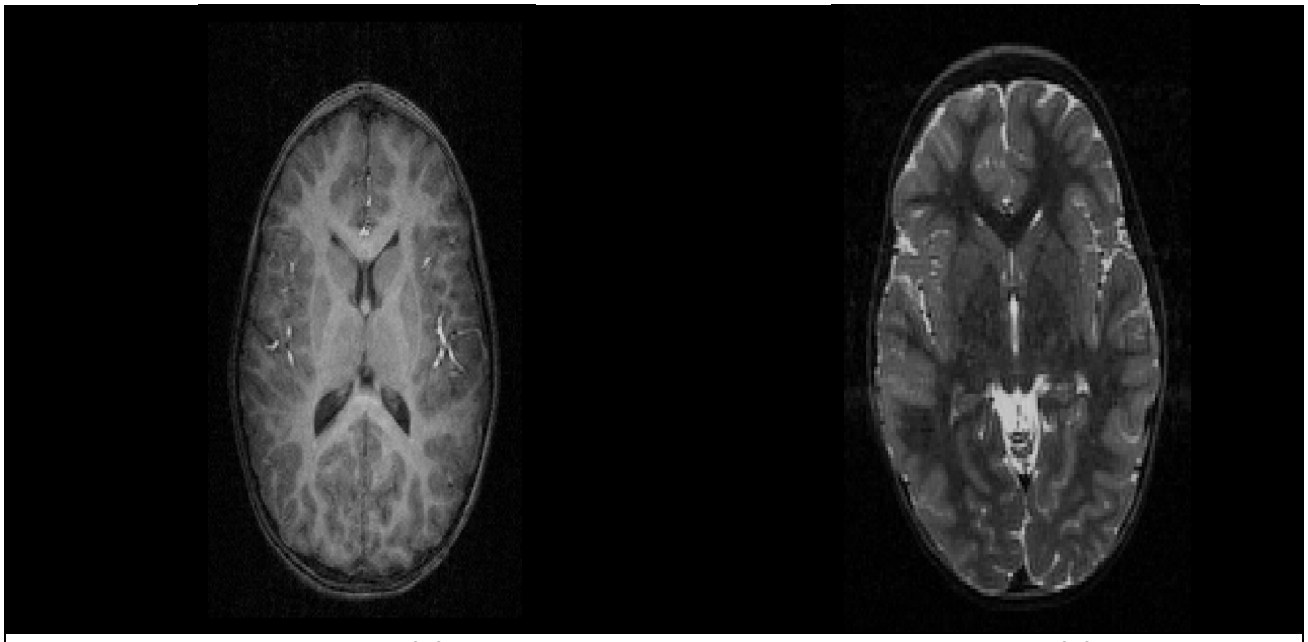
Minor (2) T1	Minor (2) T2

2 – Minor category. Small amounts of artifacts confined to a slice or two. Images are still superior quality but contain minor defects over 1 or 2 slices.



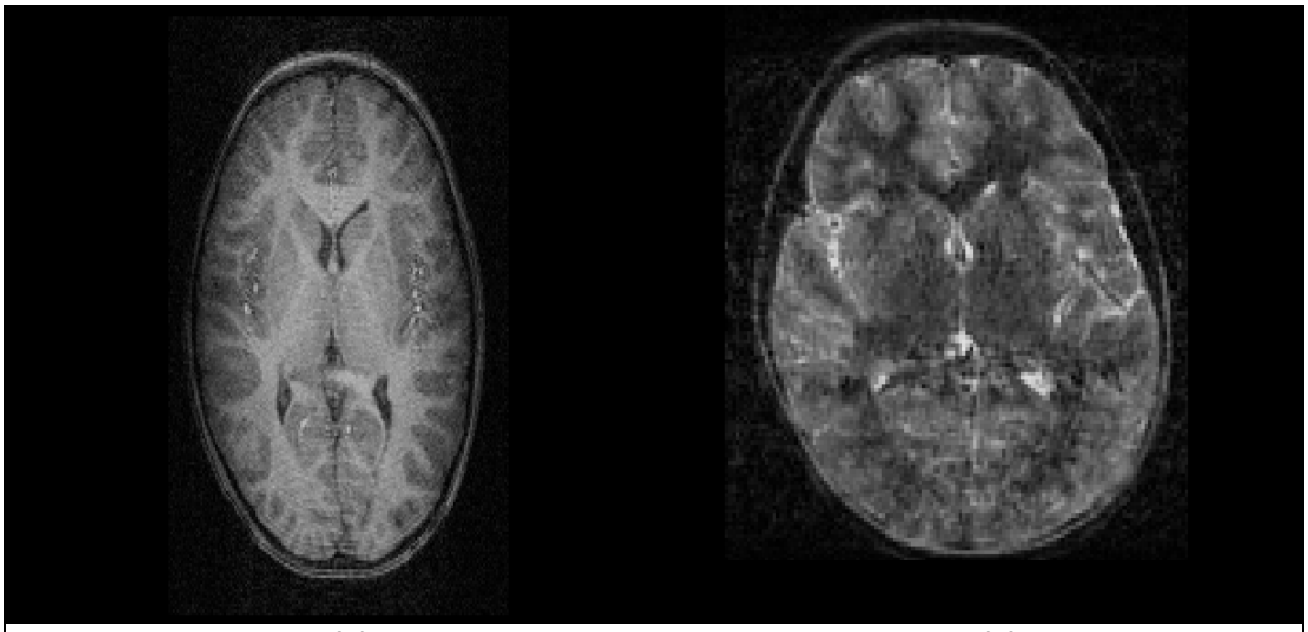
Usable (3) T1 **Usable (3) T2**

3 – Usable category. Artifacts shown here are more severe and cover more slices than in images given a rating of 2. This category still indicates a very usable image which contains moderate defects over a few (about 1 – 2) slices or minor defects over several slices.



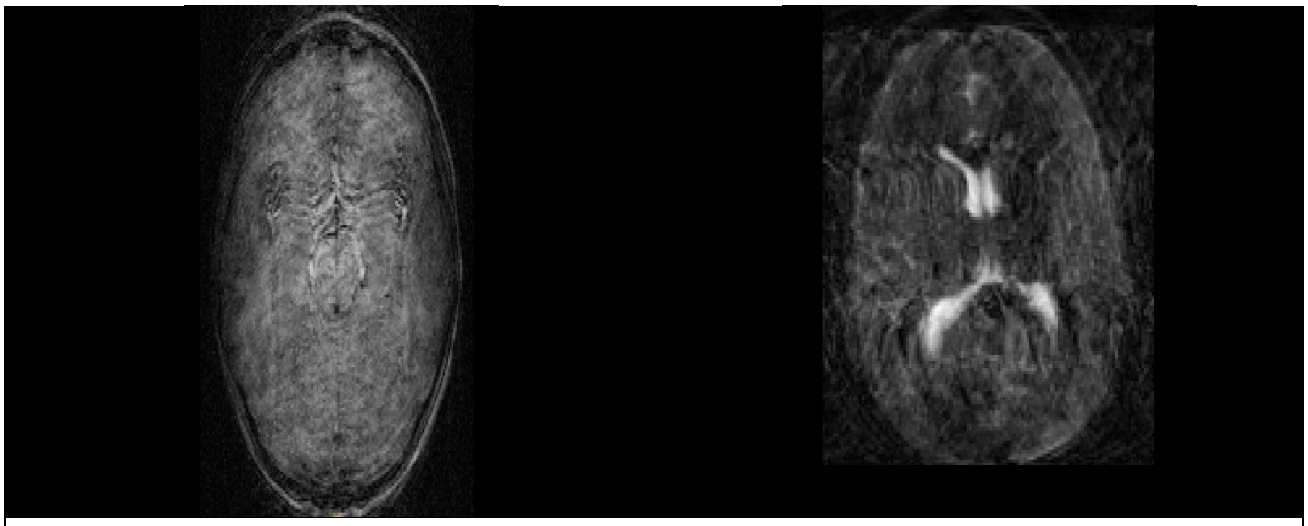
Conditional (4) T1 **Conditional (4) T2**

4 – Conditional category. The images above show moderate artifacting on many slices. Processing images in this category may possibly produce unreliable results. However, images in this category may be included or not, based on severity of artifacts. The conditional category indicates heavy artifacts on a slice or two or moderate artifacts across several slices.



Poor (5) T1	Poor (5) T2

5 – Poor category. Images in the poor category are typically excluded, however some may be salvageable. This category contains images with heavy artifacts on a slice or two or moderate artifacts throughout. The T1 image represents the moderate artifacting throughout that fits the category and could potentially be salvaged. The T2 represents the heavy artifacting that would be found on multiple slices, and would be excluded.



Bad (6) T1	Bad (6) T2

6 – Bad category. Images fitting the bad category contain heavy artifacts throughout and should be excluded from any further processing. These images show heavy artifacting throughout.

IV. Overview of Motion Artifact QC: 4-point Scale

The accurate classification of grey matter, white matter and cerebrospinal fluid is depended on the quality of MR images. Motion can cause concentric bands of high intensity on the MR images resulting in poor contrast of the brain tissues. The CONTE and TWINS studies employ a four scale quality rating system to determine the suitability of MRI images and to qualitatively control for motion artifact in statistical analyses. There is a 10% overlap of images that each technician will rate.

Inter and Intra-reliability test

Prior to completing the motion artifact QC, each rater will complete an inter/intra-reliability test to ensure consistency between and within raters. This is analyzed using the intraclass correlation and the interclass Pearson correlation (Hallgren, 2012).

4-point Scale

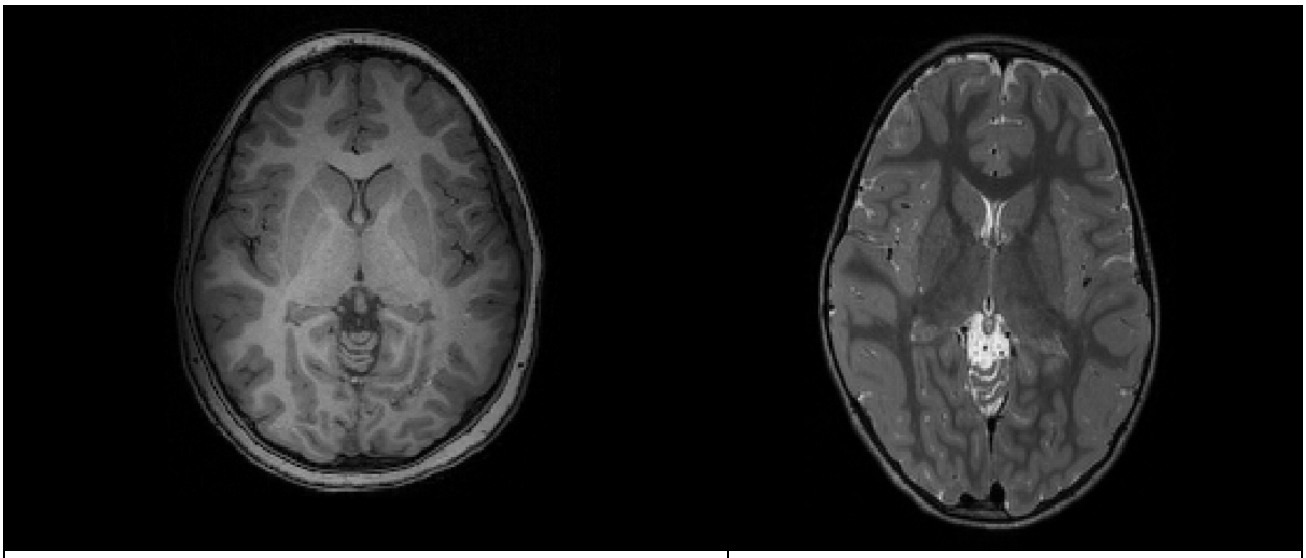
Raters employ the 4-point scale to rate the motion of each image. Increasing numbers correspond to increasing motion artifacts. The best possible score is 1, where there is no motion noted in any slices of the MR images. The worse score is 4 where the motion is great and the image would likely be considered as unusable data. The severity of concentric banding (ringing) will determine the quality of each image. This scale is based on the Blumenthal et al. (2002) method and has been used for a previous study in our lab (Lyall et al., 2014).

The following table outlines how each rating aligns with artifact severity.

Rating	Meaning	Artifacting	Location/Distribution
1	Good	No visible motion	throughout the image
		Minor artificating	Minor banding on many slices
2	Mild	Moderate artifacting	Hemodynamic and eye motion artifact noted in many ventral slices
		Mild artifacting	on many slices throughout the image
3	Moderate	Heavy artifacting	on more than two slices but not throughout
		Moderate artifacting	throughout the image
4	Severe	Heavy artifacting	throughout the image

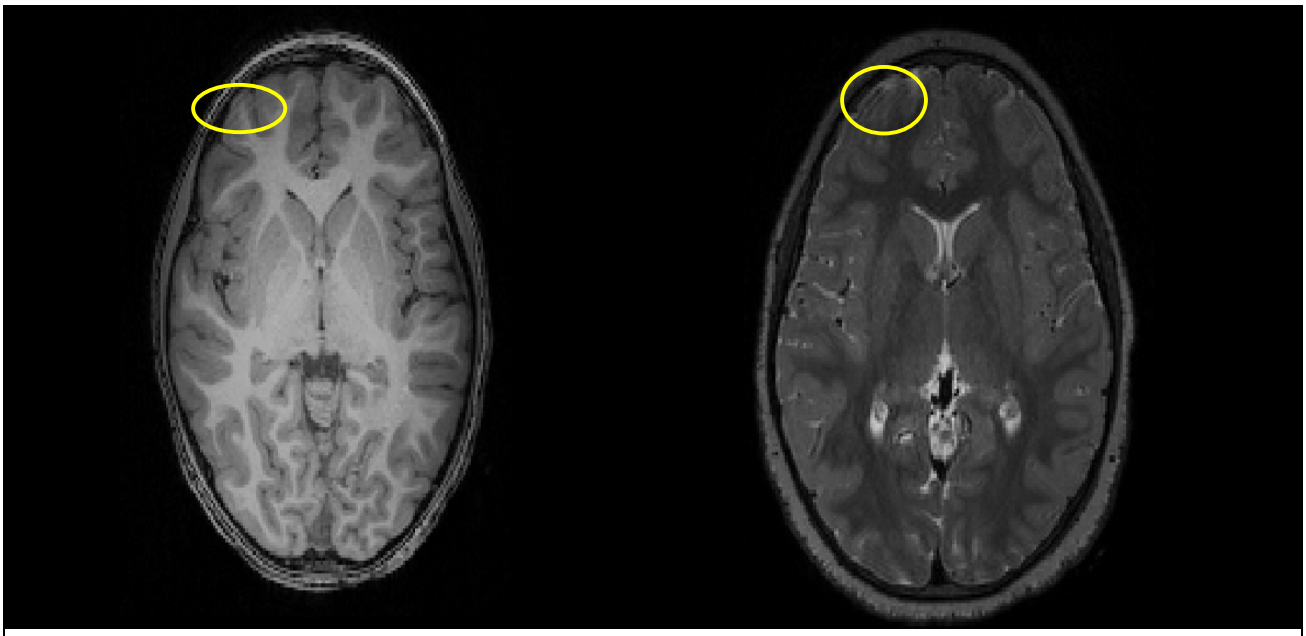
III. Image Motion Quality Examples

The following screenshots illustrate the 4-point rating scale for both T1 and T2 images. While these images are viewed and evaluated in all three dimensions, the screenshots are shown in the axial plane for simplicity.



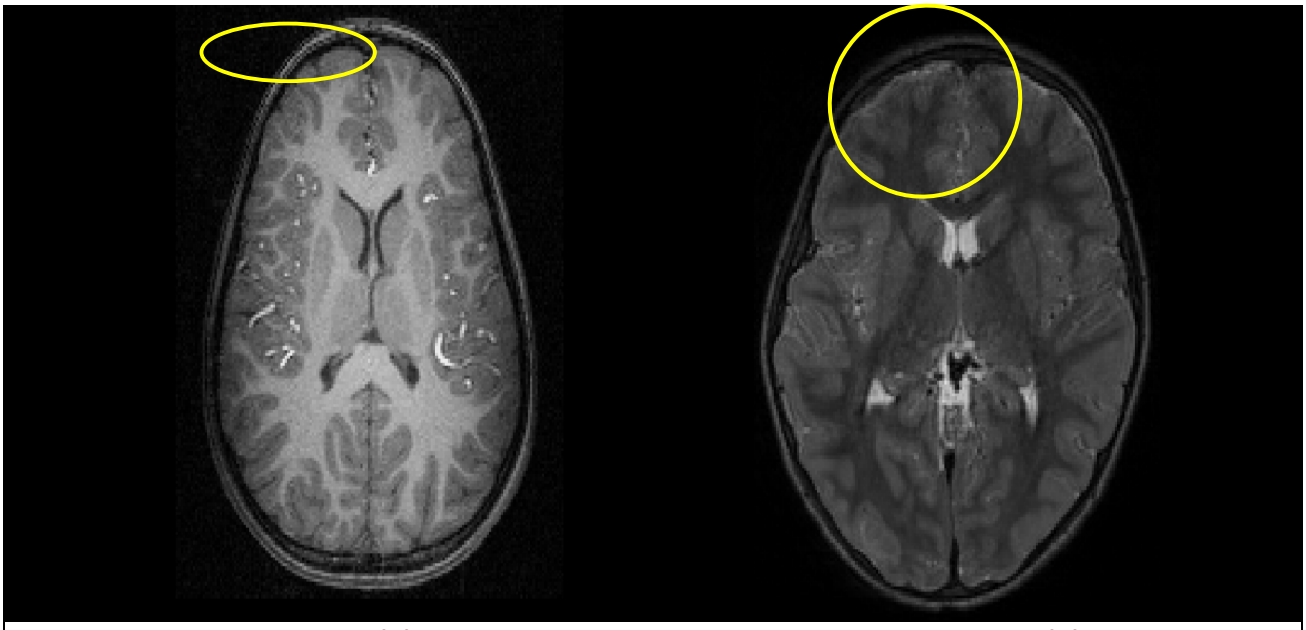
Good (1) T1	Good (1) T2

1 – **Good category.**
No artifacts detected throughout image. These images are free from any visually detectable defect.



Good (1) T1	Good (1) T2

1 – Good category.
Minor- Small amounts of artifacts confined to a slice or two. Images are still superior quality but contain minor defects.

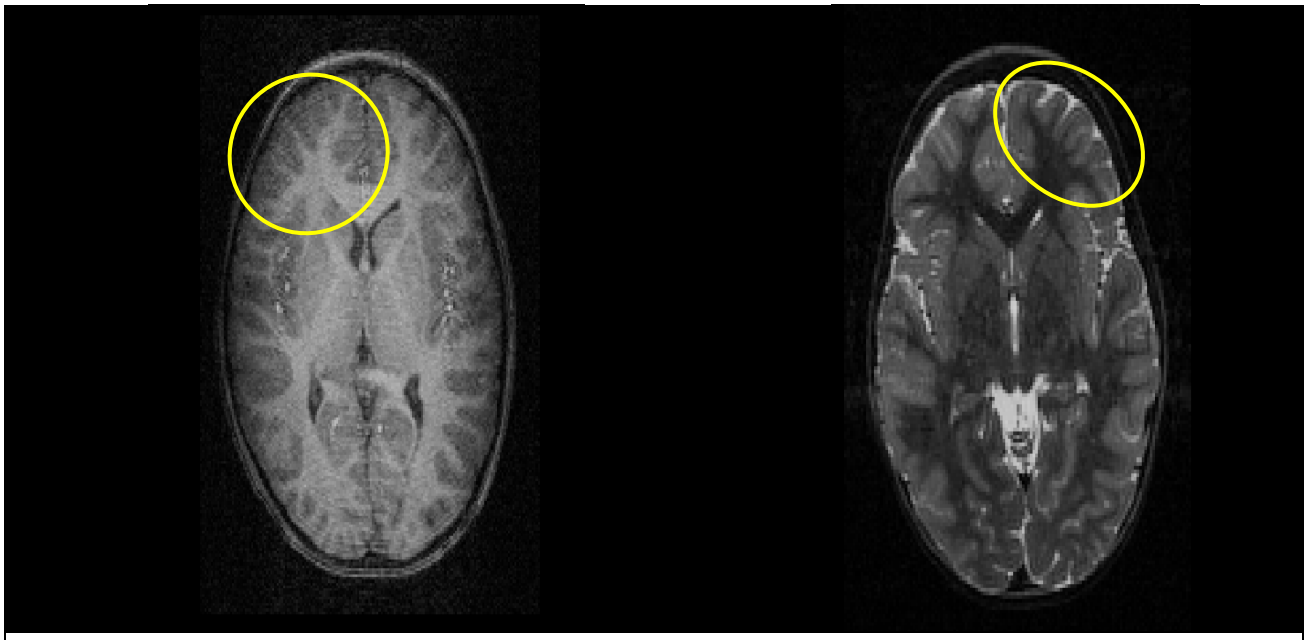


Usable (2) T1	Usable (2)

2 – Mild category.
Mild: Motion artifacts shown here are more severe and cover more slices than in images given a rating of 1. This category still indicates a very usable image which contains moderate defects over a few slices or minor defects over several slices.

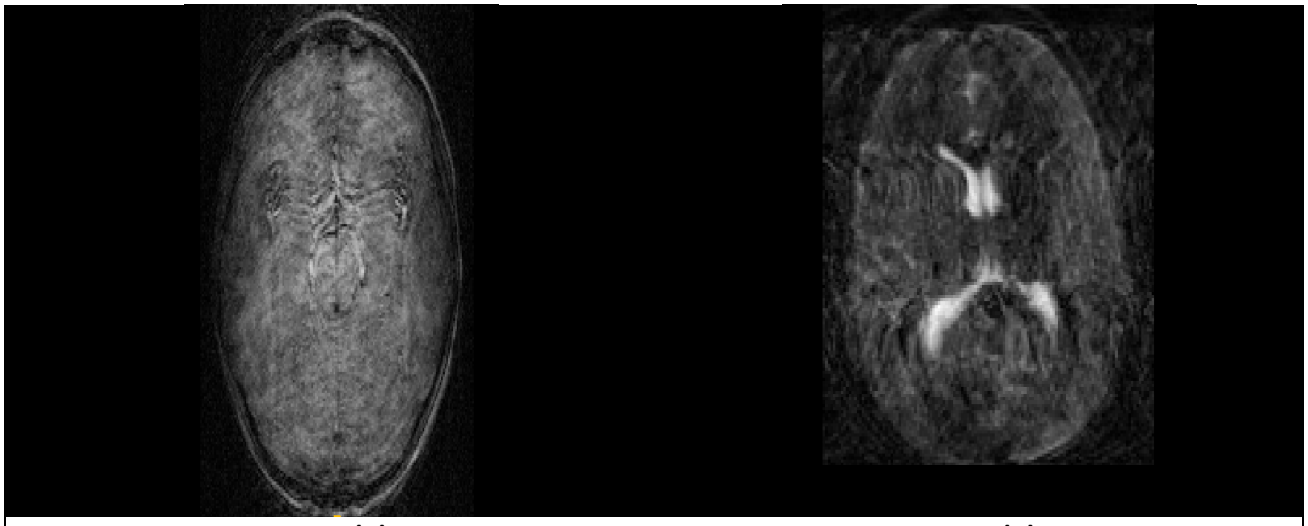
Mild (2) T1	Mild (2) T2

Mild: Hemodynamic or eye motion in many slices, however there is very little to no banding noted in other slices.



Conditional (3) T1	Conditional (3) T2

3 – Moderate category.
The images above show moderate motion on many slices. Processing images in this category may possibly produce unreliable results. The moderate category indicates heavy artifacts on a slice or two or moderate artifacts across several slices.



Severe (4) T1	Severe (4) T2

4 – Severe category.
Images fitting the severe category contain heavy motion artifacts throughout and possibly should be excluded from any further processing.

Figure of motion artifact rating from Blumethal et al. (2002)(Blumenthal, Zijdenbos, Molloy, & Giedd, 2002):

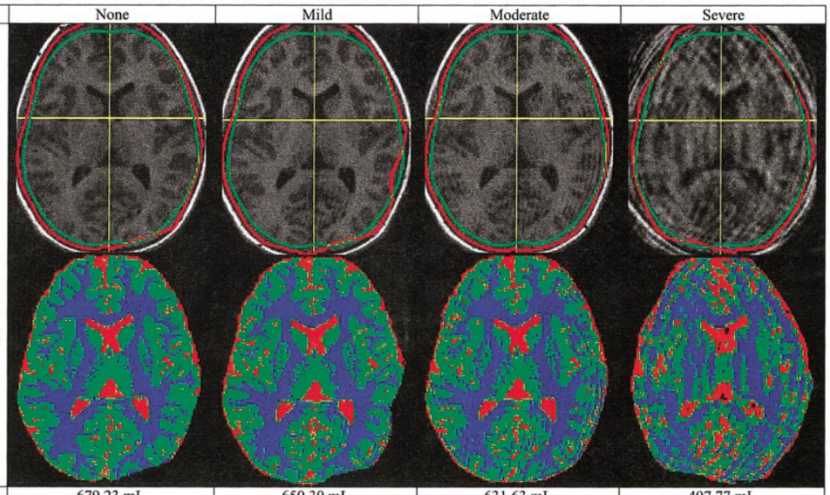
Motion	None	Mild	Moderate	Severe
Example MRI from the same subject Green=GM Blue=WM Red=CSF				
Total GM Volume for above example	679.23 mL	650.30 mL	631.63 mL	497.77 mL
Total GM Volume for sample	727.04 ± 75.19	714.27 ± 70.44	704.83 ± 29.33	582.62 ± 141.75

FIG. 1. Total GM volume for individual case and for the entire sample.

In-between Slices Motion

In addition to these common artifacts, some images have “in between slices” motion, where there is a visible change in the position of the brain when scrolling through slices. This motion should also be considered in the overall Motion Score.

V. Terminology

ARTIFACT –

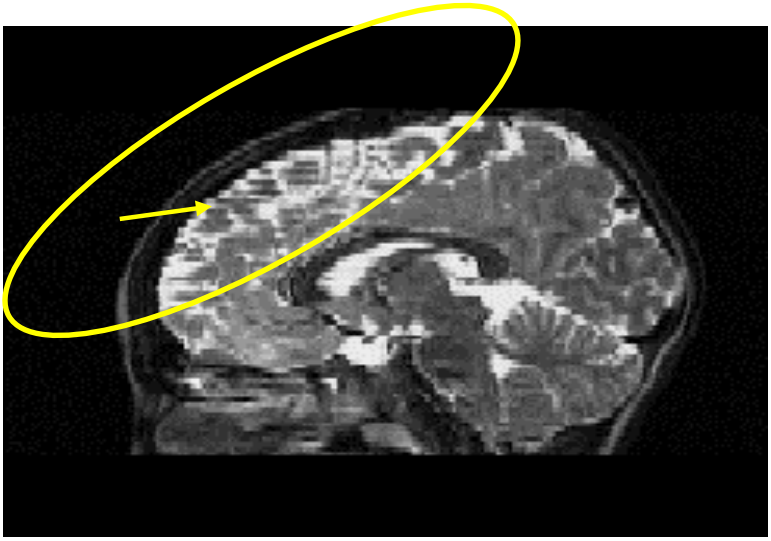
For purposes of this document, an artifact is an imperfection in the image derived from an MRI scan, which may or may not interfere with processing MRI scan data. The artifacts rated in this quality control system are usually created through patient movement within the scanner or a natural process in the human body.

MOTION –

Movement of the patient's head within the scanner is known as "motion". Sample images of motion artifacts follow.

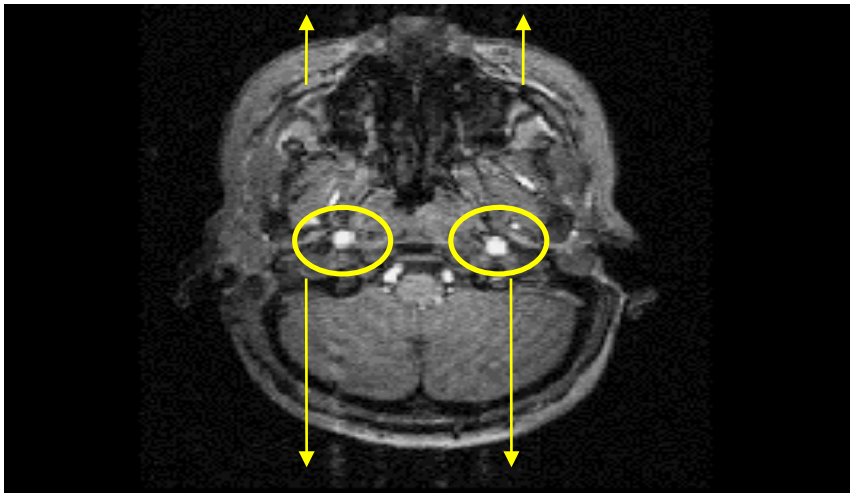
Patient head motion commonly creates artifacts that render images unusable. In this example, a T1 image shows such severe artifacts that further analysis would not be performed

Patient head motion commonly creates a "venetian blind" artifact in interleaved images, such as in this T2 weighted image. The venetian blind artifact is most evident in the sagittal plane.



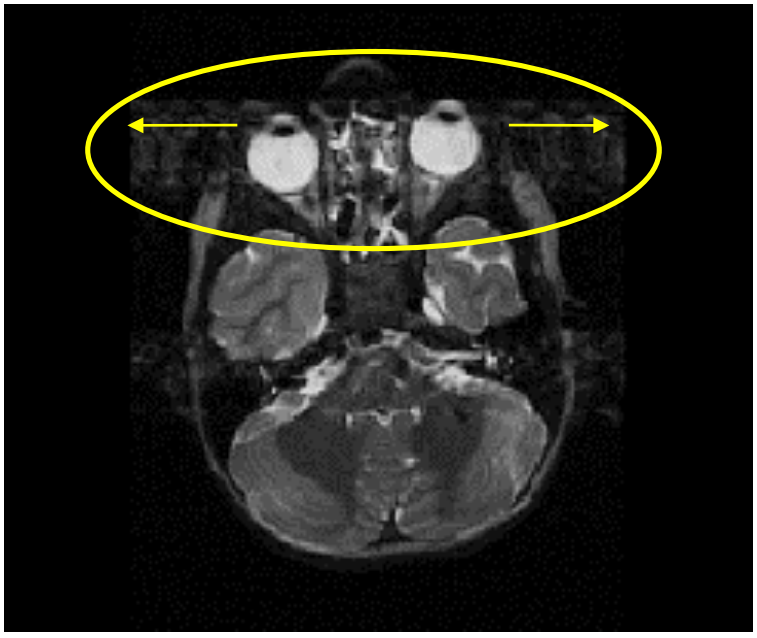
HEMODYNAMIC –

Natural processes in the human body, such as blood flow, can create artifacts in an MR image. This artifact can be easily identified by examining blood vessels in the axial plane. A sample image of a hemodynamic (blood flow) artifact follows.



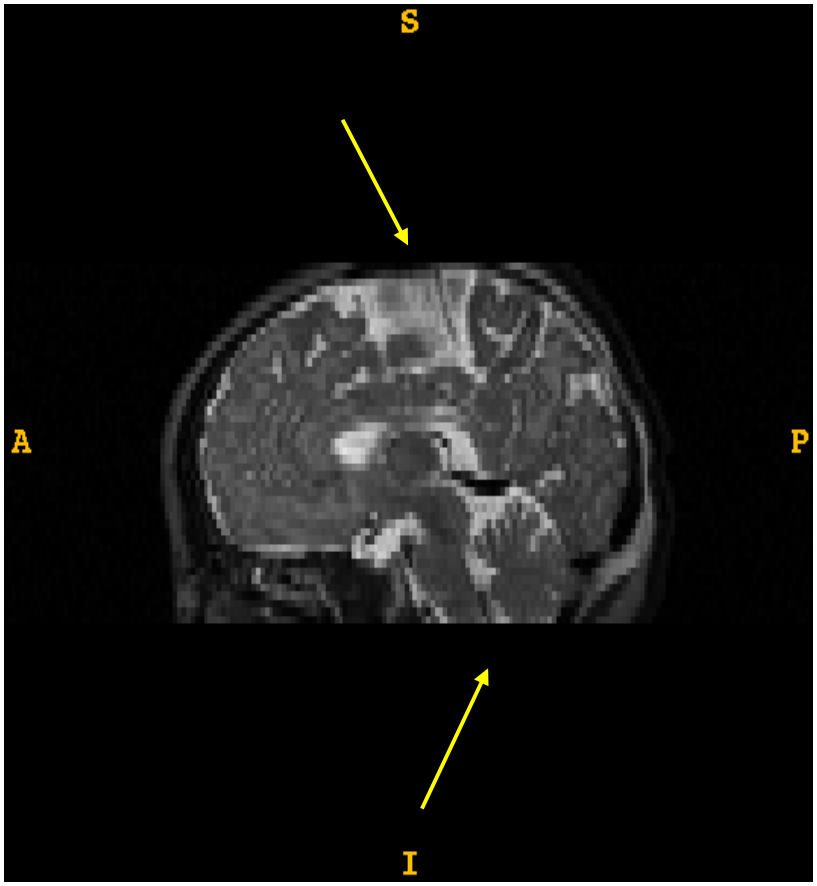
EYE MOTION –

Patient eye movement during scan causes eye motion artifacts. This artifact is most evident in the axial plane.



COVERAGE –

A cropped image occurs when the scan fails to fully cover the brain and part of the brain is excluded from the image. In this example, note the severely cropped cerebellum and missing superior slices.



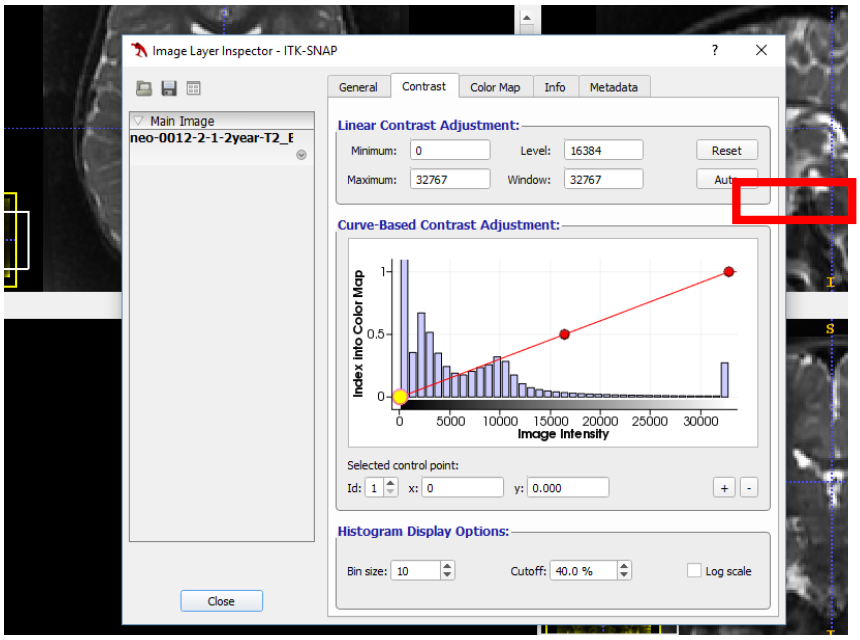
SLICE –

A slice is any 1 aspect of a 3-D planar view of an image as presented by the viewing software. A slice can be viewed in the sagittal, coronal or axial aspect in most visualization software.

VI. Instructions

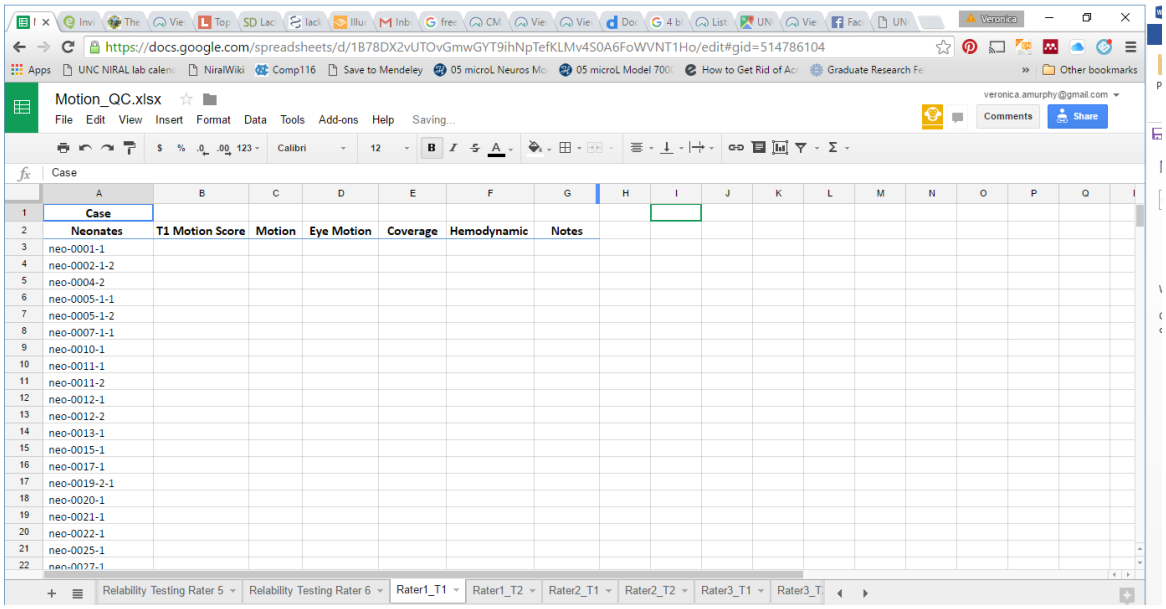
***Some Tips:

The intensity of the images can be changed using Ctrl + I. You can then press Auto and it will automatically readjust the contrast. The **Introduction** gives a brief description on how to use InsightSnap.



VII. Recording the Motion Artifact Rating

Motion quality control ratings are tracked using a Google Doc spreadsheet for each timepoint. A "Stats" sheet is also included to track the percentage of images that have been controlled. The first column contains the, Motion Score- will range from 1-4. The last four columns track the presence (1) or absence (0) of artifacts by the type of artifact in each image, (motion, eye motion, coverage, hemodynamic). Each sheet will have a column for notes where raters can discuss any issues or if there was noted in between motion.



References

Blumenthal, J. D., Zijdenbos, A., Molloy, E., & Giedd, J. N. (2002). Motion Artifact in Magnetic Resonance Imaging: Implications for Automated Analysis. *NeuroImage*, *16*(1), 89–92. http://doi.org/10.1006/nimg.2002.1076

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22833776

Lyall, A. E., Shi, F., Geng, X., Woolson, S., Li, G., Wang, L., … Gilmore, J. H. (2014). Dynamic Development of Regional Cortical Thickness and Surface Area in Early Childhood. *Cerebral Cortex (New York, N.Y. : 1991)*, 1–9. http://doi.org/10.1093/cercor/bhu027

March 2018
GreyLevel QC Rating