HHS Public Access

Author manuscript

Ultrason Imaging. Author manuscript; available in PMC 2024 December 05.

Published in final edited form as:

Ultrason Imaging. 2024 May ; 46(3): 151–163. doi:10.1177/01617346241236160.

Human observer sensitivity to temporal noise during B-mode ultrasound scanning: characterization and imaging implications

Matthew T. Huber[1], **Katelyn M. Flint**[1], **Patricia J. McNally**[2], **Sarah C. Ellestad**[3], **Gregg E. Trahey**[1,4]

[1]Department of Biomedical Engineering, Duke University, Durham, NC, USA

[2]Department of Women's and Children's Services, Duke University Hospital, Durham, NC, USA

[3]Division of Maternal-Fetal Medicine, Duke University School of Medicine, Durham, NC, USA

[4]Department of Radiology, Duke University School of Medicine, Durham, NC, USA

Abstract

This work measures temporal signal-to-noise ratio (SNR) thresholds that indicate when random noise during ultrasound scanning becomes imperceptible to expert human observers. Visible noise compromises image quality and can potentially lead to non-diagnostic scans. Noise can arise from both stable acoustic sources (clutter) or randomly varying electronic sources (temporal noise). Extensive engineering effort has focused on decreasing noise in both of these categories. In this work, an observer study with five practicing sonographers was performed to assess sonographer sensitivity to temporal noise in ultrasound cine clips. Understanding the conditions where temporal noise is no longer visible during ultrasound imaging can inform engineering efforts seeking to minimize the impact this noise has on image quality. The sonographers were presented with paired temporal noise-free and noise-added simulated speckle cine clips and asked to select the noise-added clips. The degree of motion in the imaging target was found to have a significant effect on the SNR levels where noise was perceived, while changing imaging frequency had little impact. At realistic *in vivo* motion levels, temporal noise was not perceived in cine clips at and above 28 dB SNR. In a case study presented here, the potential of adaptive intensity adjustment based on this noise perception threshold is validated in a fetal imaging scenario. This study demonstrates how noise perception thresholds can be applied to help design or tune ultrasound systems for different imaging tasks and noise conditions.

Keywords

Temporal SNR; image quality; noise sensitivity; acoustic exposure; adaptive imaging; fetal ultrasound

Corresponding author: Matthew T. Huber, Department of Biomedical Engineering, Duke University, Room 1427, FCIEMAS, 101 Science Drive, Box 90281, Durham, NC 27708, USA, matthew.huber@duke.edu.

Introduction

Ultrasound B-mode image quality has improved dramatically since the clinical introduction of this modality in the mid-20th century. These advances can be attributed in large part to engineering efforts improving the signal-to-noise ratio (SNR) of the data collected and displayed by ultrasound systems.[1,2] The signal in the SNR calculation is the echo returning from the target of interest, while several potential noise sources can comprise the SNR denominator.[3] One category of noise is referred to as “clutter” noise. Clutter noise sources arise from acoustic propagation phenomena, such as wavefront aberration, multiple scattering events, or reflections from bright off-axis targets.[4–6] In contrast, temporal noise sources arise from thermal effects in the system circuitry or from internal or external electromagnetic sources corrupting the electrical information in the ultrasound system.[7] These temporal noise sources are well-modeled by Gaussian noise profiles, distorting the underlying acoustic signal level. Both clutter and temporal noise sources can decrease feature detectability and target contrast, so minimizing the visual impacts of these noise sources is essential for high quality imaging results.

Completely eliminating ultrasound noise sources is impossible. Often, a more realistic engineering goal is to achieve sufficiently high SNR so any noise that is present minimally impairs the system performance. Generally, this quality evaluation requires sonographer input to provide a subjective opinion of whether noise levels are acceptable. Unfortunately, soliciting sonographer input at every development stage in a research environment or during live scanning is not feasible. If sonographer noise sensitivity were quantified, the resultant SNR thresholds could convey when temporal noise is at acceptable or unnoticeable levels, aiding in the system development or parameter adjustment process.

Adjustments to reduce the impact of temporal noise involve trade-offs. Once noise has been mitigated to the point where its effects are no longer visible, further noise reduction reinforces the negative aspects of these trade-offs. One method for minimizing noise in ultrasound system design is use of internal components with little inherent thermal noise. Additionally, electronic shielding to minimize interference can improve quality.[8] However, such choices may increase the cost, size, or weight of the system. Another option is temporal averaging, often adjusted through a “persistence” setting, which reduces the temporal variance of image pixels. However, with the moving targets typically encountered during *in vivo* imaging, blurring and decreased resolution are observed.[9] The imaging frequency can also be reduced to decrease signal attenuation, but resolution is again decreased with the use of longer wavelength pulses. Amplifying tissue signal magnitude by using longer or higher-pressure pulses is another method for increasing SNR. This has the drawback of increasing Mechanical Index (MI) or Thermal Index (TI) output levels that assess the potential for damaging cavitation or tissue heating effects.[10] Even below maximum output limits, regulatory bodies recommend the ALARA (As Low As Reasonably Achievable) principle be followed to keep output as low as reasonable to mitigate these potential harmful effects.[11,12] The downside associated with each method of increasing temporal SNR motivates finding the limit at which such SNR improvements are no longer helpful.

Published accounts of how different SNR levels impact clinical ultrasound imaging are lacking. To the authors' knowledge, the most relevant work was performed in a different imaging modality, X-ray fluoroscopy.[13] This paper explores the SNR where noise transitions from being visible to imperceptible. At any SNR higher than the noise perception threshold, the observer will not be able to perceive whether noise is present. The fluoroscopy work seeks to find this visibility threshold, following the premise that in conditions where noise is imperceptible, clinical readers are minimally impaired. The featured experiment of that paper showed two side-by-side clips where only one side had added noise. Observers were asked to select the noisy clip. Across observers, the average noise perceptibility threshold was approximately 27 dB. Determining the related noise perceptibility threshold in ultrasound imaging requires a revised observer study using parameters and conditions relevant for this use case.

Ultrasound is unique relative to other imaging modalities in that the tissue signal has a dominant speckle pattern due to echo interference from scatterers within the imaging point spread function.[14] Evaluating temporal noise perceptibility in a speckle field, rather than a target with constant magnitude, is necessary to make any result relevant for ultrasound. Ultrasound images are also commonly compressed and displayed on a limited dynamic range given the large differences in signal intensity across an image. Logarithmic compression of the information with gray scale color mapping over a set dynamic range is employed for the noise perceptibility study undertaken here to match common scanner standards. Different structures within the field of view of an ultrasound image have scatterers of different magnitude, creating regions of varying SNR. For the ultrasound perceptibility study, constant magnitude scatterers were used in the simulated field of view to better capture the observer sensitivity at a specific SNR. Extension of the constant magnitude study results to natural images requires consideration of regional SNR variation. Each region of an image could be compared to a noise perceptibility threshold to determine whether noise would be visible in those areas. Considering pure ultrasound speckle for the noise perception task makes the result applicable to a wide range of imaging scenarios. While this paper does not specifically vary beamforming methods, attenuation, speed-of-sound levels, or simulate spatial variation in the point-spread function, imaging with these different conditions will still yield speckle signals. The central question of how perceptible noise is in the presence of that speckle signal is largely unchanged.

Given the basic speckle target visualized over the selected dynamic range, two factors expected to impact the perceptibility threshold are the degree of target motion and the imaging frequency. *In vivo* imaging targets move involuntarily due to physiological factors. As speckle-generating scatterers move, the speckle signal decorrelates.[15] We hypothesize visual fluctuations caused by motion-induced speckle decorrelation will mask the presence of noise and lead to decreased observer noise sensitivity compared to cases with less motion. In ultrasound, acoustic frequency can dramatically affect SNR, resolution, and ultimately the imaging performance. Especially in the presence of axial motion, higher frequencies experience greater speckle decorrelation as the imaging point spread function is more constrained.[16] We hypothesize higher frequencies will thus have lower noise perceptibility thresholds, following the same masking effect expected with motion. Evaluating varied motion and frequency conditions gives insight into the factors influencing temporal noise

visibility, and provides an indication of whether noise is expected to be visible or not in certain imaging conditions.

Noise perceptibility insights are valuable in a variety of system optimization aims. A particularly well-suited example is in adaptive adjustment of acoustic intensity. Despite the recommendations to follow the ALARA principle, studies suggest sonographers rarely consider acoustic output during clinical scanning given the numerous other tasks they are required to perform.[17,18] As such, the default output levels on a system are likely the settings most routinely used in practice. While these default levels may vary by the manufacturer and system configuration, in the authors' experience MI settings between 0.7 – 1.2 are typically observed. The authors are not aware of any systems that adapt these output levels in response to image quality or ALARA considerations. Recent efforts seeking to automatically adjust intensity following ALARA recommendations evaluated image quality as intensity was swept, using the point at which a spatial coherence-based image quality metric begins to asymptote as a guide for selecting scan intensity.[19,20] However, that approach would not necessarily achieve consistent visual imaging characteristics as both clutter and temporal noise sources impact the spatial coherence. The temporal noise perceptibility threshold could instead be used as a standard for recommending acoustic output following the ALARA principle. Adjusting output to always achieve SNR matching the temporal noise perceptibility threshold would result in consistent visual image quality conditions. This approach was retrospectively performed with previously collected data from a fetal imaging study[20] to give insight into potential acoustic output reductions that would result from adopting this method.

This paper expands on a preliminary investigation of ultrasound temporal noise perceptibility.[21] This expanded text updates the SNR recommendations provided in that initial study by reexamining the analysis process and aligning it with standard psychophysics methods. Additional exploration of the results and discussion of the study context are also provided.

Methods

The approach employed here for determining temporal noise human observer perceptibility thresholds is based on the psychophysics method of constant stimuli paired with a two-alternative forced choice (2AFC) experiment. This combined experimental approach is prevalent in studies evaluating perception levels, which relates to this work's goal of establishing temporal noise perception thresholds in ultrasound systems. Under the method of constant stimuli, observers are asked to identify whether a stimulus is present in various example cases, with the stimulus level varying randomly between those cases.[22] The observer sensitivity threshold is evaluated by performing numerous repeat observations at different stimulus levels. This constant stimulus method is combined with the 2AFC method, presenting paired options to the observer. Only one of the paired options has an added stimulus, which the observer is tasked with identifying.[23] A standard criterion in psychophysics research for determining observer stimulus sensitivity is correct identification of that stimulus 50% of the time, absent the effect of random guessing. Because random guessing in the two-alternative framework would lead to correct identification 50% of the

time, correctly identifying the noise in 75% of trials will be the threshold defining observer noise perceptibility.[22]

The simulated ultrasound clips presented to the observers in this study have been designed to capture specifics of an ultrasound imaging environment without containing extraneous information that could confound the noise threshold definitions. For accurate representations of motion, the effect of transducer and physiological movements were captured from *in vivo* acquisitions. These motion profiles were used in simulation, shifting speckle-generating point targets to create cine clips of moving speckle fields. Temporal noise was added to the cine clips to vary the SNR. To determine noise perceptibility thresholds for ultrasound imaging, the noise-free and noise-added clips were presented side-by-side to human observers to assess their ability to select the clip with added noise. Fig. 1 presents a visual representation of the experimental process and evaluation methods.

A. Clinical Motion Profile Generation

Profiles of the motion present during ultrasound scanning were obtained by tracking target movement during scans acquired for a prior clinical study.[24] The data were collected with an approved Duke Health Institutional Review Board (IRB) protocol from the same eighteen volunteers as in that prior study, but with a 5-second-long acquisition period and an 11 frames per second data collection rate. For these acquisitions, the sonographer was asked to hold the transducer stationary over target structures in the fetal abdomen or cranium. Volunteers were not asked to hold their breath during the acquisitions, as this is typically not requested during clinical fetal scans. This simulates the condition of a sonographer saving data for a clinical acquisition, the point in a scan where having minimal noise is essential for proper diagnoses. A Siemens Acuson SC2000 ultrasound scanner and a 6C1 transducer (Siemens Healthcare, Issaquah, WA, USA) were used for pulse-inversion harmonic imaging. The transmit frequency was set to 2 MHz, and the images were formed from receive data with a center frequency of 4 MHz. In all acquisitions, axial and lateral translation could be recognized to varying degrees and tracked with the 2D ultrasound data. Some acquisitions also had observable elevation motion, recognized by targets changing in shape or size as new cross-sections of the target were imaged. These cases with observed elevation motion were excluded, resulting in the 29 motion profiles used in the simulation study.

Motion during these clinical acquisitions was tracked using MATLAB’s image registration function, imregdemons. This function utilizes a diffusion model for image-to-image matching based on pushing points in the current image into alignment with the reference image.[25,26] The first frame of the cine clip was compared to each subsequent frame to estimate the cumulative displacement throughout the acquisition. The parameters used for the image registration function were 25-iterations of the algorithm with a Gaussian smoothing factor of 7.0 applied between iterations as a motion profile regularization parameter. This iteration count and regularization parameter were selected based on empirical assessment of the correspondence between the acquired clinical cine clip and a recreated clip produced by displacing the first frame of the cine capture with the acquired motion profile.

Motion was tracked throughout the field of view imaged in each acquisition. Imaging depths and lateral ranges varied slightly between acquisitions, so only a subset region that was present in all acquisitions was considered. This region was a sector centered at a depth of 7 cm, where the lateral extent and axial span were 3 cm. The displacement magnitude, velocity, and temporal behavior of the motion were analyzed in this region to characterize the level of involuntary motion present during fetal scanning.

B. Speckle and Noise Cine Clip Simulation

Scattering fields were generated by randomly distributing uniform magnitude scatterers through an 11 cm by 11 cm region, with a density of 15 scatterers per resolution cell. A resolution cell is defined as the area of the axial by lateral full-width at half maximum for the imaging point-spread function (PSF). Using 15 scatterers per resolution cell ensures fully developed speckle.[27] Motion was introduced by shifting the location of scatterers following the clinical motion profiles.

Three different motion levels were simulated. The first case had no motion (referred to here as a 0x motion level). The simulated scattering targets remained at fixed locations throughout the cine clip, representing an extreme case of no transducer or physiologic motion during a 5-second imaging period. The second case replicated a more realistic motion level expected when a sonographer has located a target and is doing their best to keep the view steady (1x motion). This case was simulated by applying the motion profiles developed in the previous methods section to generate displaced scatterer locations for each frame of the eventual cine clip. Additionally, exaggerated motion levels (3x motion) were simulated where the normal steady-imaging displacements were increased by a factor of three before being applied to the scatterer field.

With the scatterer positions determined for each frame of the cine clip, speckle maps were created by convolving a Field II generated PSF with the scattering field.[28,29] Speckle maps at the three motion levels were created for a 2 MHz frequency imaging system with a F/2 aperture and 0.75 fractional bandwidth. Additionally, 4 MHz frequency was also used to create simulated clips for comparison at the standard (1x) motion level. To limit the sonographer observation task to approximately 1-hour in length, 0x and 3x motion levels were not simulated at 4 MHz.

Noise was introduced by first initiating a random white noise profile for each frame of the speckle cine clip. These random noise fields were then matched to the spatial frequencies of the imaging system by multiplying the 2D Fourier transformed PSF by the frequency content of the white noise, before performing the inverse Fourier transform on the resultant frequency information. The noise frame power was scaled relative to the speckle signal power to create the specified temporal SNR level for each simulated clip. The SNR levels simulated were 20 dB, 25 dB, 27.5 dB, 30 dB, 32.5 dB, 35 dB, and 40 dB. The noise frames generating these SNR levels were added to the beamformed speckle frame prior to envelope detection to create the noise-added cine clips. Each noise-free and noise-added cine clip was cropped to a sector with an axial span of 3 cm and a lateral width of 3 cm across the center.

C. Human Observer Study Process

Videos with side-by-side noise-free and noise-added clips were created with the left-right positioning of each clip randomized. For each clip, the pixel values underwent logarithmic compression and were displayed on a 60 dB dynamic range. A series of videos was created for presentation to each observer in the study. Each video series featured 16 paired clips at each of the SNR levels, motion levels, and frequencies. With seven SNR levels, three motion levels at 2-MHz, and one motion level at 4-MHz, a total of 448 cine clips comprised each observer video series. For each paired cine clip presented to an observer, their task was to select which one they perceived to be corrupted by random noise.

Five observers individually viewed the study clips. Each observer was a sonographer with a specialty in either echocardiography or fetal imaging. Observers ranged in experience from between 4 to 32 years of clinical practice. Study clips were viewed in a darkened room, similar to the lighting conditions ideally used in clinical settings to make details more visible on the display. The monitor used was 39 cm along its diagonal and presented the images on the 1920 by 1080 pixel display with a 60 Hz refresh rate and 8-bit grayscale color depth. Before viewing the 448-clip study set, observers were trained in the task by looking through a standard set of 10–15 example clips. Once the study began, observers were presented with each clip in randomized order. Each clip played once for its 5-second duration and the observer was asked to select the noisy clip. This clip length and study design were chosen so as to give observers time to form their impressions of the clips and make a judgement which had noise, while encouraging them to perform these evaluations in a real-time manner. Observers were allowed to take breaks from viewing clips as needed to prevent eye fatigue. Each participant completed their observation task in approximately 1 hour.

The proportion of correct noise identifications at each SNR level was computed for each observer. Linear interpolation between the noise levels in the study was performed to determine the SNR level at which the 75% threshold level is crossed. While this interpolated value gives the most likely SNR threshold based on the study results, significance testing was performed to better quantify these results. A one-sample proportion z-test evaluated the observer results with a null hypothesis H_0: $p = 0.75$ and alternative hypothesis H_a: $p \neq 0.75$. For the combined observer results where N=80 with a p-value of 0.05, proportions between 0.65 and 0.85 would fail to reject the null hypothesis, while observed proportions below or above this level would lead to the null hypothesis being rejected in favor of the alternative. With the assumption that perception decreases monotonically with increasing SNR, the significance test limits were paired with the observer results to place bounds on the SNR levels that could potentially meet the 75% threshold defining the perception limit.

D. Adaptive Intensity Application

Data collected for a previous adaptive intensity study[20] were reexamined to recommend acoustic output MI levels required to reach different temporal SNR thresholds. These data were collected during fetal imaging in 35 second-trimester volunteers with healthy, singleton pregnancies. Three acquisitions were performed with each volunteer, removing the transducer from the volunteer between acquisitions to collect non-identical views. In each acquisition, data at 16 MI levels from 0.15 MI to 1.20 MI were collected in a 2 cm

axial by 4 degree lateral region of interest located in the fetal abdomen. The center of the axial span was located at the transmit focal depth, which was chosen based on the target depth. Pulse-inversion harmonic imaging pulses were transmitted twice on each of the 5 lines in the region of interest, resulting in four consecutive acoustic transmissions along each line. To calculate temporal SNR, the coherence between each of the two repeated lines of harmonic data was calculated. This temporal coherence (ρ_t) of the summed data is related to the temporal SNR (SNR_t) with the following equation[30]:

$$SNR_t = \frac{\rho_t}{1 - \rho_t}$$

Interpolating between the median temporal SNR measured in the region at each of the 16 varied MI acquisitions, the MI expected to yield temporal SNR conditions ranging from 15 dB to 40 dB was determined. Across three acquisitions in each of the 35 volunteers, the mean MI required to reach each of these temporal SNR targets as well as the distribution of those MIs were evaluated for each perceptibility threshold. This information can be used to recommend acoustic output settings during fetal imaging that achieve conditions with imperceptible temporal noise. As scanning takes place and temporal SNR is calculated and updated, the recommended MI values could be automatically selected by the system to enable adaptive real-time intensity adjustment.

Results

A. Clinical Motion Profiles

Considerable variability in velocity was observed during the 5-second scan period of each acquisition. Fig. 2 shows distributions representative of minimum, mean, and maximum frame-to-frame velocities observed during each scan. Motion in the target region typically ranged from 0.3 mm/s at its slowest up to 3 mm/s, while exhibiting an average velocity around 1 mm/s. The ranges of each boxplot are almost entirely separate, demonstrating that velocity was highly variable throughout a scan when comparing the lowest to highest velocity timepoints.

Supplemental gifs 1, 2, and 3 show clinical cine clips for three acquisitions representing cases with relatively low, medium, or high motion levels, respectively. Along with the clinical clip on the left of the gif, the center frame of each gif includes a recreation of the clinical capture generated by translating the first frame of each clip by the measured motion profile. In the far-right frame, the relative distance and direction of displacement from the original position are shown. While the simulated motion clip does not have the overlaid temporal noise of the clinical cine clip acquisition, the structures underlying that noise appear to move similarly between the two cases. Achieving this visual match between the motion in the clinical clip and the simulated recreation informed the selection of the parameters for the motion-tracking tool.

The motion observed in each clip could potentially arise from sources such as breathing, maternal or fetal heartbeats, other body movements, or involuntary transducer translation. Fig. 3 shows the median axial and lateral displacements in the field of view in each of

the supplemental sample gifs previously presented. In instances where displacement occurs, most of that observed motion appears cyclical, especially the axial displacements in the case with the median velocity level across volunteers and in the high velocity case. The cyclical nature of the displacements points towards breathing or a heartbeat as primary factors causing motion, since both these sources would create motion in regular intervals.

B. Human Observer Study Results

After introducing the clinical motion to the simulated speckle clips and adding varying amounts of noise, the paired video clips were presented to each observer. As an example of the types of clips and noise levels shown to the observers, supplemental gif 4 is provided, which combines the 20 dB, 27.5 dB and 35 dB SNR clips for a 1x motion level profile, with the corresponding noise free speckle clip. In addition, supplemental gif 5 shows examples of the different motion levels and frequencies at the 20 dB SNR level. After presenting clips in a pair-wise manner to the reviewers, the proportion of correct identifications of the noisy video clips was calculated for each participant. Fig. 4 displays the average proportion correct at the different SNR and motion levels when simulating 2 MHz imaging. The error bars around average values represent the standard deviation of correct identification proportion among the five observers. Unique relationships between observer sensitivity to noise and SNR exist with each motion level. In the absence of motion (0x motion), observers are highly sensitive to motion until around 32.5 dB. Above that level of temporal SNR, the ability of the observers to detect the noise drops off sharply. At the standard clinical motion levels (1x motion) a more gradual and nearly linear decrease is observed, from nearly 100% correct identification of the noisy clip at 20 dB down to only 55% correct identification at 40 dB. At the highest motion level (3x motion) the observers appear largely insensitive to noise as even at the lowest SNR the noisy clip was correctly identified on average at a 65% rate.

In Fig. 5, the clinical 1x motion level results for the 2 MHz imaging case are shown with the addition of results for the 4 MHz imaging case. Despite the different frequencies, the trend in each of these curves is largely similar. On average, the noise was slightly less perceptible at each SNR with the higher frequency.

The 75% threshold for defining the perceptibility threshold was applied to the results in Fig. 4 and 5. Table I lists the SNR at the perceptibility cutoff for each case based on linearly interpolating between measured SNR levels and rounding to the nearest integer. These SNR perception cutoff levels are approximations of the underlying true perception level as a result of experimental randomness. Following a one-sample proportion z-test, the last column of Table I describes the SNR levels that define the possible range for the true perception threshold. All SNR levels in the experiment outside the listed range were associated with proportions that had a statistically significant difference ($p < 0.05$) from the 75% perception level. Additional observer evaluation of SNR clips within this range could be performed to define the SNR perception threshold more accurately. As hypothesized and observed in Fig. 4, increasing motion is associated with noise becoming imperceptible at lower SNR.

C. Adaptive Intensity Application

Measured temporal SNR levels in three of the 16 MI levels of the fetal abdomen acoustic output sweeps are shown in Fig. 6. The three repeat acquisitions performed in each volunteer at each output level results in three markers of each color. The 35 volunteers are ordered based on the average measured SNR of the 0.5 MI acquisitions. The low- and medium-output show good SNR consistency between repeats, considering the probe was removed and replaced with each acquisition. At 1.2 MI, there is higher variability, suggesting that at higher SNR levels, the result is increasingly sensitive to probe placement and motion during the measurement.

Interpolating between the measured SNR associated with each of the known output levels in the fetal imaging data, the expected acoustic output required to yield a given SNR condition is estimated. The expected average, standard deviation spread, and range for the acoustic output required to reach a range of SNR conditions are displayed in Fig. 7. As the target SNR rises, the output necessary to reach those imaging conditions increases. Depending on the SNR threshold, the average MI required might be as low as 0.3 or as high as 1.1. With higher average MI, the standard deviation is also slightly increased. There is always a large difference between the minimum and maximum boundaries of the plot, indicating a wide range in the MI potentially required to reach a given SNR.

The MI sweep from 0.15 to 1.2 used in this experiment constrains the possible MI values to that measured range. In all cases, at least 0.2 MI was required to reach 15 dB, the lowest SNR shown in Fig. 7. When trying to achieve SNR conditions of 27 dB or higher, the maximum MI of 1.2 was not always sufficient. Fig. 8 shows the percent of acquisitions where the threshold SNR could not be achieved. To consistently achieve SNR of 30 dB or above, more than 1.2 MI would be required.

These results indicate which acoustic output conditions would be used if MI were adaptively selected to achieve SNR conditions corresponding to imperceptible noise. Table II lists the average and standard deviation in output achieving each of the SNR thresholds, along with the range of output that would be recommended to meet those conditions.

Discussion

A. Clinical Motion Profiles

The cine clips providing the basis for the motion profiles were acquired by an expert sonographer with more than 25 years of scanning experience. This experience likely contributes to the seemingly low displacement and velocity observed during the acquisitions. Little motion was observed originating from a shifting transducer, indicating the sonographer was able to keep the probe stationary over the target. In addition, the sonographer controlled volunteer positioning, acoustic window selection, and data acquisition timing, so sub-optimal views with high anatomical motion were likely avoided. A less experienced sonographer may not hold the probe in position as well or may scan with conditions that would lead to higher motion levels.

Obtaining motion profiles representative of the cine clips was confounded by two factors: the presence of large hypoechoic regions in the field of view and, at- times, high levels of temporal variation in the signals. The hypoechoic regions resulted from amniotic fluid surrounding the fetus. With little signal in these areas, correlation-based speckle tracking failed to produce accurate motion profiles, leading to the choice of the ìmregdemons' image registration method. However, temporal noise had a tendency to confound this method as its temporally uncorrelated nature caused frame-to-frame variation unrelated to the underlying motion. The initial result was substantial high-frequency quivering in the motion profiles not representative of the target motion. While the baseline noise associated with these data was not collected, an essential task for the motion tracking algorithm is to ensure that noise is not leaking through to the generated motion profiles. With use of the large regularization parameter, sufficient smoothing occurred to mitigate observed high-frequency disturbances resulting from that noise. While this smoothing increases confidence that observed motion originates from the target and not the noise, the smoothing likely leads to a mild underestimation of the true motion.

With both sonographer experience and motion smoothing in the image registration method, the simulated motion level is likely subtly less than the motion that might be encountered during a typical scanning session. As motion has an impact on the temporal noise perceptibility level, this bias could lead the SNR thresholds for noise perceptibility reported here to be slightly higher than would otherwise be required. This tendency towards slightly higher SNR would increase ALARA MIs in the intensity adjustment scenario, but it would also increase certainty temporal noise has been mitigated in the scans to the point where it is not detrimental to image quality.

B. Human Observer Determined Temporal SNR Perceptibility

The observation that temporal noise perceptibility is closely related to motion level supports the hypothesis that motion masks the presence of temporal noise. The increased speckle decorrelation with larger motion levels potentially contributes to this masking effect. Because motion levels are dynamic during ultrasound scanning, this indicates the temporal noise perceptibility levels can change with time. When sonographers are sweeping the transducer and looking for a structure of interest, the motion level would be higher than when that target is located and the transducer is held steady. While it may be task dependent, the steady imaging case (1x) standard motion level noise perceptibility threshold is likely the most relevant for clinical applications. All of the higher-motion time periods would still have imperceptible noise if the steady-motion limit were used given their lower noise perception threshold. Using the steady motion SNR threshold would lead to visible noise in the artificial case where there is no (0x) motion, but this condition is impractical during *in vivo* scanning.

While axial and lateral translations were represented in the clinical motion profiles generated, elevation motion could not be tracked. Clinical acquisitions with minimal observed elevation translation were selected for this reason. However, elevation motion does occur in practice, and would result in additional speckle decorrelation relative to cases where there is no elevation motion. Because speckle decorrelation can be visually indistinguishable

from noise, we can hypothesize elevation motion would result in lower temporal noise perceptibility SNR thresholds, following similar trends as Fig. 4 where more motion leads to lower perception thresholds.

An additional tested hypothesis related to speckle decorrelation was that using higher frequency data would lead to a lower perceptibility threshold. As seen in Fig. 5, for each SNR, the 4 MHz imaging case has, on average, lower correct identification of the noisy clip. This follows the hypothesized result, but the degree of separation and the resultant shift in the perception threshold is minor. It is possible that with a more extreme frequency difference the effect on the perception threshold would be greater. However, because doubling the frequency from 2 MHz to 4 MHz did not substantially impact the noise perception level, frequency does not appear to be strongly tied to the noise perceptibility level. As a result, the temporal SNR perception levels determined here are applicable to imaging cases with relatively similar frequencies.

When an ultrasound imaging transmit frequency is selected, an important associated consideration is whether echoes received at the fundamental or harmonic frequencies will be used for image formation. For both imaging modes, the same SNR thresholds are expected to apply. In either fundamental or harmonic imaging, the acoustic pulse interacts with the target scattering medium, producing a speckle interference pattern. For such a speckle pattern, regardless of the mechanism used to generate it, the overlaid temporal noise would become imperceptible at the SNR levels determined in this study. Certain processing methods can distort speckle signals. For example, averaging with a large enough sampling kernel smooths over the brightness variations of the speckle, producing an image with more consistent brightness. Such processing steps may, in turn, result in different SNR perception levels. Future observer studies could explore to what extent speckle-reduction steps alter these perception levels.

In the simulated video clips used to define the noise perception thresholds only constant-magnitude scatterers were represented. In realistic scanning environments, clutter noise is also present in the echo signal. However, as clutter noise arises from acoustic propagation phenomena, it results in additional speckle overlaid on the tissue signal with similar temporal stability to that signal. While patterns indicating what might be clutter artifacts become recognizable to sonographers, the similar spatial and temporal characteristics of clutter and tissue signals make them otherwise visually indistinguishable. As such, when extending the noise perceptibility threshold results to cluttered environments, it is expected the combined tissue and clutter power relative to the temporal noise power comprise the relevant SNR level for comparison with the threshold SNR cutoffs found here. Future studies could test this hypothesis directly, as well as consider noise perception in hypoechoic or anechoic regions, such as in the ventricles of the heart or amniotic fluid surrounding a fetus, where clutter signal may be dominant.

During *in vivo* scanning, with both tissue echogenicity and clutter levels varying across a given imaging field of view, SNR is expected to vary throughout the visualized region. As a result, the temporal noise visibility will vary spatially as well. In general, noise will be easier to see in darker regions as the local signal power, and thus the SNR, is lower.

This phenomenon can often be seen with increased temporal noise visibility in anechoic regions or at greater imaging depths where the signal has attenuated. For all regions with SNR at or above the relevant noise perception threshold, no visual temporal noise is expected. As a result, when optimizing settings and trying to achieve conditions where temporal noise is imperceptible, defining a region of interest for consideration is essential. If the region includes a range of SNR levels, adjusting parameters so the median SNR is at the perceptibility threshold can be a target to achieve generally good image quality across the region. A more aggressive approach towards eliminating visual noise could entail considering, for example, the 25th percentile SNR in the region, and adjusting parameters until that 25th percentile SNR is at the perception threshold. Doing so would ensure 75% of the region is above that threshold value, eliminating most of the visual impacts of noise. While automatic acoustic power updates following these methods could serve as the default on a system, override measures allowing for manual control could still be provided based on user preference.

The results shown here indicate approximate perception thresholds for ultrasound imaging, but it should be noted that there are a number of factors that may impact the threshold values in other scenarios. Because these are video clips, we hypothesize the video playback conditions and quality will impact the noise perception level. Such factors might include the display update frame rate, monitor and room brightness, screen resolution, color calibration, data compression, and the dynamic range. Given these factors, the SNR levels reported here could be used as defaults when configuring an imaging setup, but potential adjustments to these levels should be considered if the imperceptible noise conditions are not observed. For systems adaptively selecting parameters based on the perception threshold, sonographers should be aware of how to adjust global perception threshold variables on their system to achieve their desired imaging performance as conditions change.

Individual sonographer adjustment of perception threshold levels could be important if sonographers were found to vary in their sensitivity to temporal noise. This study was not designed to compare performance between sonographers, but future studies could consider inter-observer variations in noise sensitivity and compare those variations to factors like experience level, training, or specialization. Intra-observer analyses would also be interesting, considering how consistent observers are throughout a study period or between repeats of the experiment done on different days. To perform such studies, more replicates with each observer would be required at each SNR level and imaging condition. Replicate counts of around 100 or more are recommended for achieving well-behaved confidence limits for 2-alternative forced choice experiments.[31] With 16 replicates at each SNR in the current study, the confidence limits for each individual observer’s results would not be suitable for making inter-observer comparisons. Instead, the average response across the five observers estimates the expected noise sensitivity in the sonographer population. These results thus serve as a suitable starting point for calibration of noise sensitivity levels that might be generally applicable, while potential follow-up studies could evaluate the required flexibility based on determining the distribution of noise sensitivity levels among the sonographer population.

C. Adaptive Intensity Application

In the clinical case study motivating the importance of quantifying noise perceptibility thresholds, the imaging target was always the fetal abdomen. Despite the consistent target, Fig. 6 shows there are widely varying temporal SNR conditions between volunteers, and even differences with repeat acquisitions in the same individual. This demonstrates the diverse range of imaging conditions encountered in a clinical setting. If acoustic output were set to a default level on a scanner, equivalent to imaging with just one of the color indicators in Fig. 6, widely varying temporal SNR would be observed across patients. In some cases the temporal SNR would well-exceed all relevant noise perceptibility thresholds, indicating unnecessary acoustic exposure. In other instances, the noise threshold would not be met, resulting in impaired image quality. In contrast, following an adaptive intensity approach, a noise perceptibility threshold could be selected to serve as a consistent visual quality level goal and the system could automatically increase or decrease acoustic output until that SNR condition is met. Acoustic power adjustments to target the desired SNR conditions could occur at a fixed period, be triggered when the system detects the probe has settled on imaging a new target, or occur when the transducer is moved to a new acoustic window.

Factors such as the imaging target and system settings will have an impact on the acoustic output required to achieve different SNR conditions, like those shown in Fig. 7. In addition, different ultrasound systems are expected to have different background temporal noise levels based on the electronic design choices made in those systems. Therefore, the MI distribution achieving a given SNR condition will vary with systems and populations. Knowing the MI distribution required to achieve a given SNR for a certain clinical task can be helpful from an engineering perspective to ensure a system is optimized for that task. However, advance knowledge of this MI distribution is not required for clinical scanning with an adaptive output scheme as the subject- and target-specific environment will inform the selected output. All that is required is calculation of the SNR at the current MI, and an adaptive adjustment to increase or decrease the scan MI in response until the desired SNR level is reached. Implemented this way in a clinical setting, the selected MI would be well-suited for a specific patient or view and that MI could be determined without needing to sweep through 16 MI levels as in this example case study.

Based on the MI levels reported in Table II, a significant advantage of following an adaptive output selection scheme is evident in the potential for reducing acoustic exposure. With standard 1x motion levels, the average MI required to reach the noise perceptibility limit was around 0.6. Default acoustic output levels are generally between 0.7 and 1.2 MI, but MI as high as 1.9 are permitted by current FDA regulations.[11] A shift to noise-perception-based output selection would immediately result in lower output than the default levels without an expected increase in visible temporal noise. This satisfies the ALARA goal in output selection. An additional advantage of reducing output includes increasing the expected runtime of battery-operated scanners.

Calculating SNR in the clinical acquisitions relies on computing the temporal coherence between sequential pulses. As part of the SNR calculation, we assume that tissue signals are coherent. However, as highlighted in Ref. 15, speckle signals decorrelate with motion. This decorrelation suppresses the calculated SNR relative to conditions where motion is

absent. As such, the true SNR values associated with the data in Fig. 7 are higher than the values represented. Therefore, the MI required to achieve a given SNR condition in practice is likely lower than the MI recommendations of that figure. To avoid the effect of motion in the SNR calculation, a null-transmit condition could be considered for evaluating background noise levels, or those noise levels could be calculated from analysis of received echo coherence in a completely stationary phantom environment.

As a result of the acoustic output constraint on the MI sweep in the clinical study, the range of possible SNR conditions was limited. At the standard clinical motion level and 2 MHz imaging condition less than 5% of acquisitions were MI limited, but if a higher SNR threshold were used, this number could increase dramatically. As a result, the expected visual quality would not be achieved. In such cases, the system would simply use the highest MI that is allowed and perhaps provide an indication to the observer that lower than desired temporal SNR is occurring. This could indicate that other adjustments could be performed to improve SNR, such as finding a better acoustic window, lowering the frequency, or increasing the persistence setting.

Conclusions

Temporal noise may be unavoidable in ultrasound imaging, but in certain imaging scenarios, this noise can be reduced to the point that it is not visible to human observers. This work determined temporal SNR threshold levels where noise becomes imperceptible. These thresholds are highly related to motion level as increased motion creates a varied signal, masking the presence of noise. Knowing quantitative signal quality levels corresponding to the noise perceptibility threshold, adaptive parameter adjustment can be performed to meet those conditions. The benefit of such an approach was demonstrated in fetal imaging, showcasing how adaptive intensity adjustment could yield consistent visual quality while, on average, decreasing acoustic output. Providing a framework for adaptive parameter adjustment promises more streamlined ultrasound procedures, as scanner operators can focus more on probe positioning and image interpretation, rather than on tweaking settings and controls. Understanding what levels of temporal noise are acceptable will also help guide the development of new ultrasound systems or techniques, as engineers can ensure the methods they develop produce relevant results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding and Acknowledgments:

This research was supported by the National Institute of Biomedical Imaging and Bioengineering through grants R01-EB017711 and R01-EB026574 and the National Science Foundation Graduate Research Fellowship Program through grant DGE 2139754. Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the National Institutes of Health or the National Science Foundation.

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Duke Health Institutional Review Board under Application No. Pro00038307.

The authors would like to thank the anonymous sonographer reviewers for participating in the observer study. Thanks also to Dr. D. Bradway, Dr. J. Long, and Dr. R. Ahmed for their help preparing the observer study.

References

1. Contreras Ortiz SH, Chiu T, Fox MD. Ultrasound image enhancement: A review. Biomed Signal Process Control. 2012 Sep 1;7(5):419–428.
2. Harvey CJ, Pilcher JM, Eckersley RJ, Blomley MJK, Cosgrove DO. Advances in Ultrasound. Clin Radiol. 2002 Mar 1;57(3):157–177. [PubMed: 11952309]
3. Vienneau EP, Ozgun KA, Byram BC. Spatiotemporal Coherence to Quantify Sources of Image Degradation in Ultrasonic Imaging. IEEE Trans Ultrason Ferroelectr Freq Control. 2022 Feb 17; 69(4):1337–1352. [PubMed: 35175919]
4. Trahey GE, Smith SW. Properties of acoustical speckle in the presence of phase aberration part I: First order statistics. Ultrason Imaging. 1988 Jan 1;10(1):12–28. [PubMed: 3291366]
5. Lediju MA, Pihl MJ, Dahl JJ, Trahey GE. Quantitative assessment of the magnitude, impact and spatial extent of ultrasonic clutter. Ultrason Imaging. 2008 Jul 1;30(3):151–68. [PubMed: 19149461]
6. Pinton GF, Trahey GE, Dahl JJ. Sources of image degradation in fundamental and harmonic ultrasound imaging using nonlinear, full-wave simulations. IEEE Trans Ultrason Ferroelectr Freq Control. 2011 Apr;58(4):754–65. [PubMed: 21507753]
7. Pettai R Noise in receiving systems. New York: Wiley; 1984.
8. Ott HW. Noise reduction techniques in electronic systems. 2nd ed. New York: Wiley; 1988.
9. Zander D, Hüske S, Hoffmann B, Cui XW, Dong Y, Lim A, et al. Ultrasound Image Optimization ("Knobology"): B-Mode. Ultrasound Int Open. 2020 Jun 1;6(1):E14–E24. [PubMed: 32885137]
10. Deng Y, Palmeri ML, Rouze NC, Trahey GE, Haystead CM, Nightingale KR. Quantifying Image Quality Improvement Using Elevated Acoustic Output in B-Mode Harmonic Imaging. Ultrasound Med Biol. 2017 Oct 1;43(10):2416–25. [PubMed: 28755792]
11. Food & Drug Administration. Marketing Clearance of Diagnostic Ultrasound Systems and Transducers | Guidance for Industry and Food and Drug Administration Staff. 2023.
12. How to Interpret the Ultrasound Output Display Standard for Diagnostic Ultrasound Devices: Version 3. J Ultrasound Med. 2019 Dec 1;38(12):3101–5. [PubMed: 31736131]
13. Luijendijk HA. Practical experiment on noise perception in noisy images. Med Imaging 1994: Image Percept. 1994 Apr 1;2166:2–8.
14. Wagner RF, Smith SW, Sandrik JM, Lopez H. Statistics of Speckle in Ultrasound B-Scans. IEEE Transactions on Sonics and Ultrasonics. 1983; 30:156–163
15. Trahey GE, Smith SW, Ramm OT von. Speckle Pattern Correlation with Lateral Aperture Translation: Experimental Results and Implications for Spatial Compounding. IEEE Trans Ultrason Ferroelectr Freq Control. 1986 May 1;33(3):257–64. [PubMed: 18291782]
16. Wagner RF, Insana MF, Smith SW. Fundamental correlation lengths of coherent speckle in medical ultrasonic images. IEEE Trans Ultrason Ferroelectr Freq Control. 1988;35(1):34–44. [PubMed: 18290126]
17. Houston LE, Allsworth J, MacOnes GA. Ultrasound is safe… right?: Resident and maternal-fetal medicine fellow knowledge regarding obstetric ultrasound safety. J Ultrasound Med. 2011 Jan 1;30(1):21–27 [PubMed: 21193701]
18. Drukker L, Droste R, Chatelain P, Noble JA, Papageorghiou AT. Safety Indices of Ultrasound: Adherence to Recommendations and Awareness During Routine Obstetric Ultrasound Scanning. Ultraschall Med. 2020 Apr 1;41(2):138–145. [PubMed: 32107757]
19. Long W, Bottenus N, Trahey GE. Lag-One Coherence as a Metric for Ultrasonic Image Quality. IEEE Trans Ultrason Ferroelectr Freq Control. 2018 Oct 1;65(10):1768–80. [PubMed: 30010556]
20. Flint K, Bottenus N, Bradway D, McNally P, Ellestad S, Trahey G. An Automated ALARA Method for Ultrasound. J Ultrasound Med. 2021 Sep 1;40(9):1863–77.
21. Huber M, Flint K, Trahey G. Human observer sensitivity to temporal noise in ultrasound imaging. In: IEEE International Ultrasonics Symposium, IUS. 2022. p. 1–4.

22. Colman A A Dictionary of Psychology. 4th ed. New York: Oxford University Press; 2015. method of constant stimuli.
23. Colman A A Dictionary of Psychology. 4th ed. New York: Oxford University Press; 2015. two-alternative forced-choice task.
24. Long W, Hyun D, Choudhury KR, Bradway D, McNally P, Boyd B, et al. Clinical Utility of Fetal Short-Lag Spatial Coherence Imaging. Ultrasound Med Biol. 2018 Apr 1;44(4):794–806. [PubMed: 29336851]
25. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal. 1998 Sep 1;2(3):243–60. [PubMed: 9873902]
26. Vercauteren T, Pennec X, Perchant A, Ayache N. Diffeomorphic demons: Efficient non-parametric image registration. Neuroimage. 2009 Mar 1;45(1):S61–72. [PubMed: 19041946]
27. Dutt V, Greenleaf JF. Speckle Analysis Using Signal to Noise Ratios Based on Fractional Order Moments. Ultrason Imaging. 1995 Oct 1;17(4):251–68. [PubMed: 8677561]
28. Jensen JA. Field: A program for simulating ultrasound systems. In: 10th Nordicbaltic Conference on Biomedical Imaging. 1996. p. 351–353.
29. Jensen JA, Svendsen NB. Calculation of Pressure Fields from Arbitrarily Shaped, Apodized, and Excited Ultrasound Transducers. IEEE Trans Ultrason Ferroelectr Freq Control. 1992 Mar 1;39(2):262–267. [PubMed: 18263145]
30. Friemel BH, Bohs LN, Nightingale KR, Trahey GE. Speckle decorrelation due to two-dimensional flow gradients. IEEE Trans Ultrason Ferroelectr Freq Control. 1998 Mar 1;45(2):317–27. [PubMed: 18244183]
31. Mckee SP, Klein SA. Statistical properties of forced-choice psychometric functions: Implications of probit analysis. Percept Psychophys. 1985;37(4):286–98. [PubMed: 4034345]

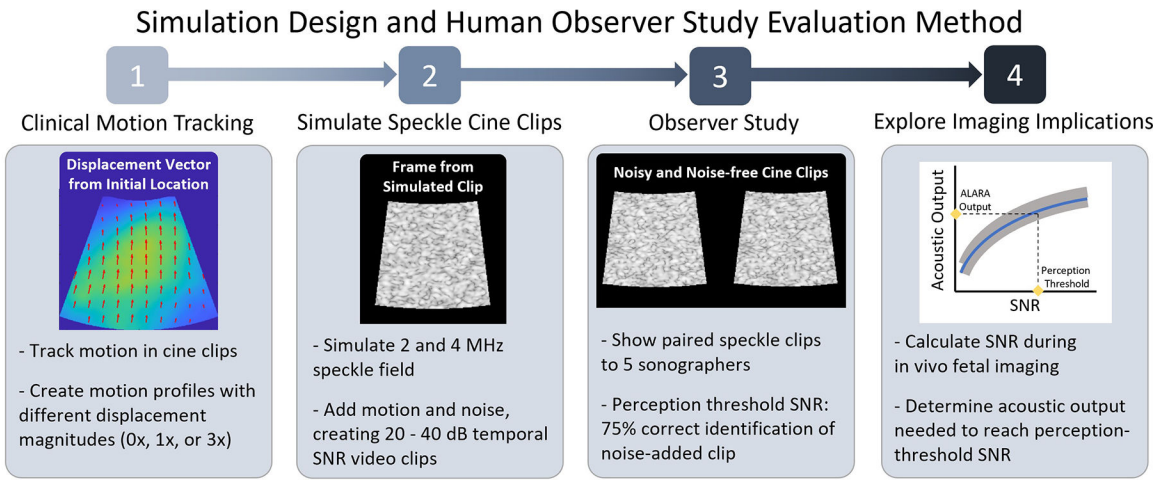

Fig. 1:

Overview of methods followed to prepare for and perform temporal noise perceptibility human observer study. Following study, evaluation of a potential use-case for the perception thresholds was performed in an adaptive intensity application with fetal imaging data.

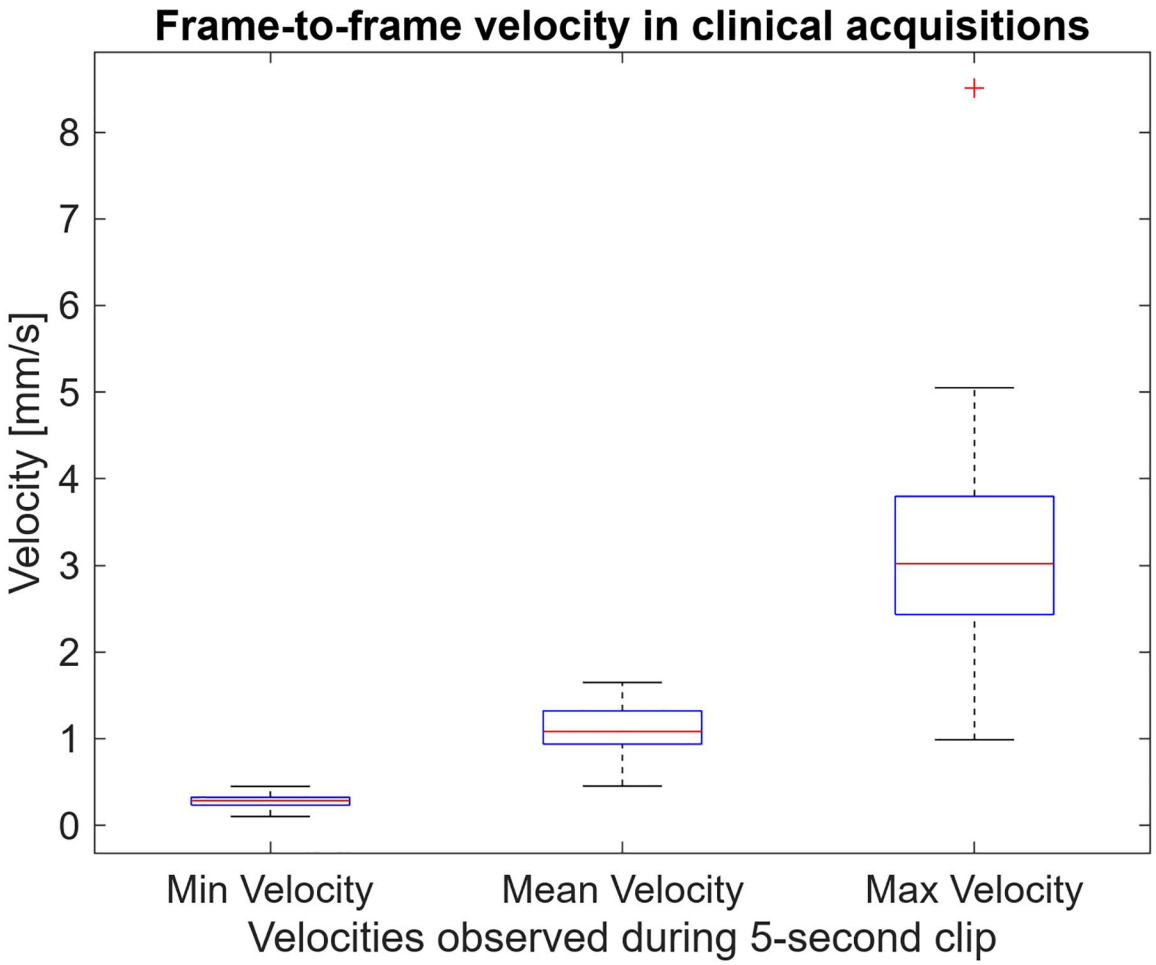

Fig. 2:

Minimum, maximum, and average frame-to-frame velocities observed in 5-second long clinical acquisition cine clips. The complete separation between min and max velocities shows all targets had periods where they were nearly stationary and other moments of increased motion.

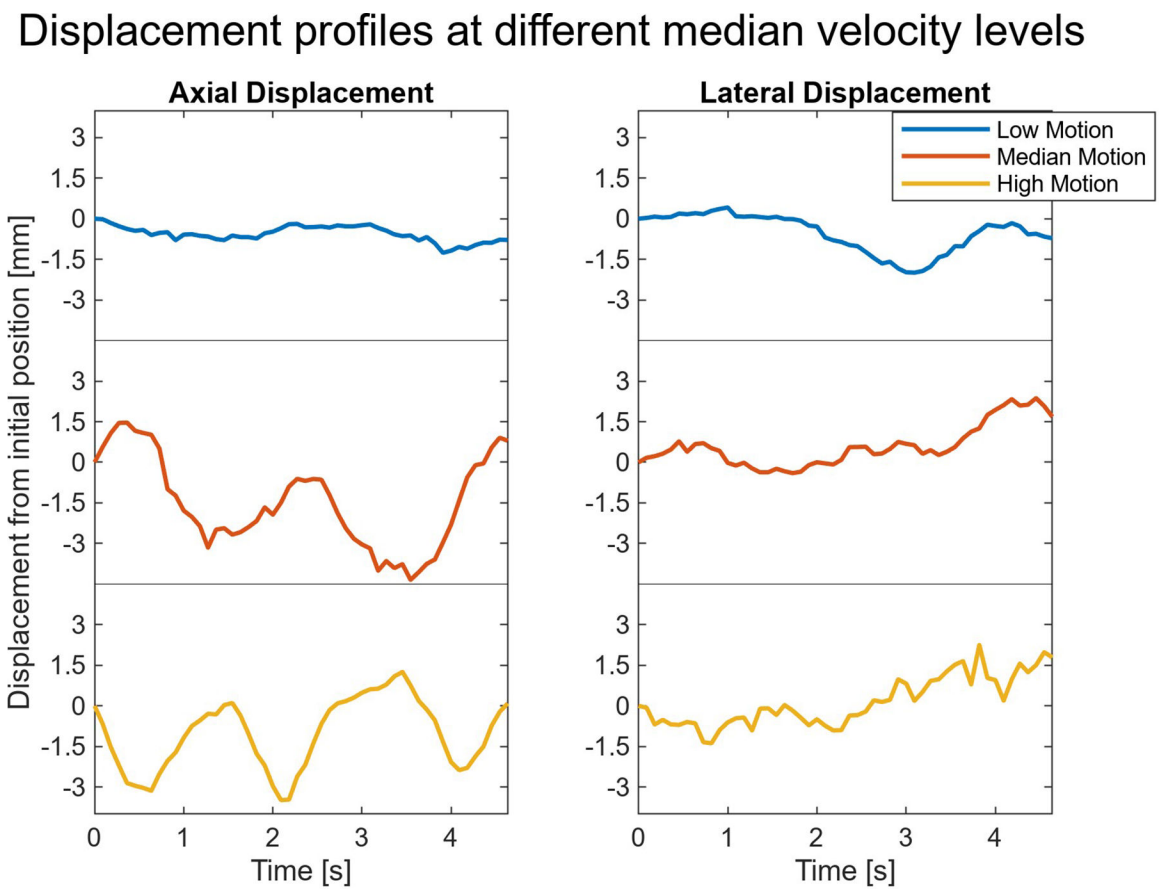

Fig. 3:

Displacement observed through time in the three supplementary sample clips. Most of the motion in the scans appears cyclical, suggesting heartbeats or breathing are the primary causes of motion.

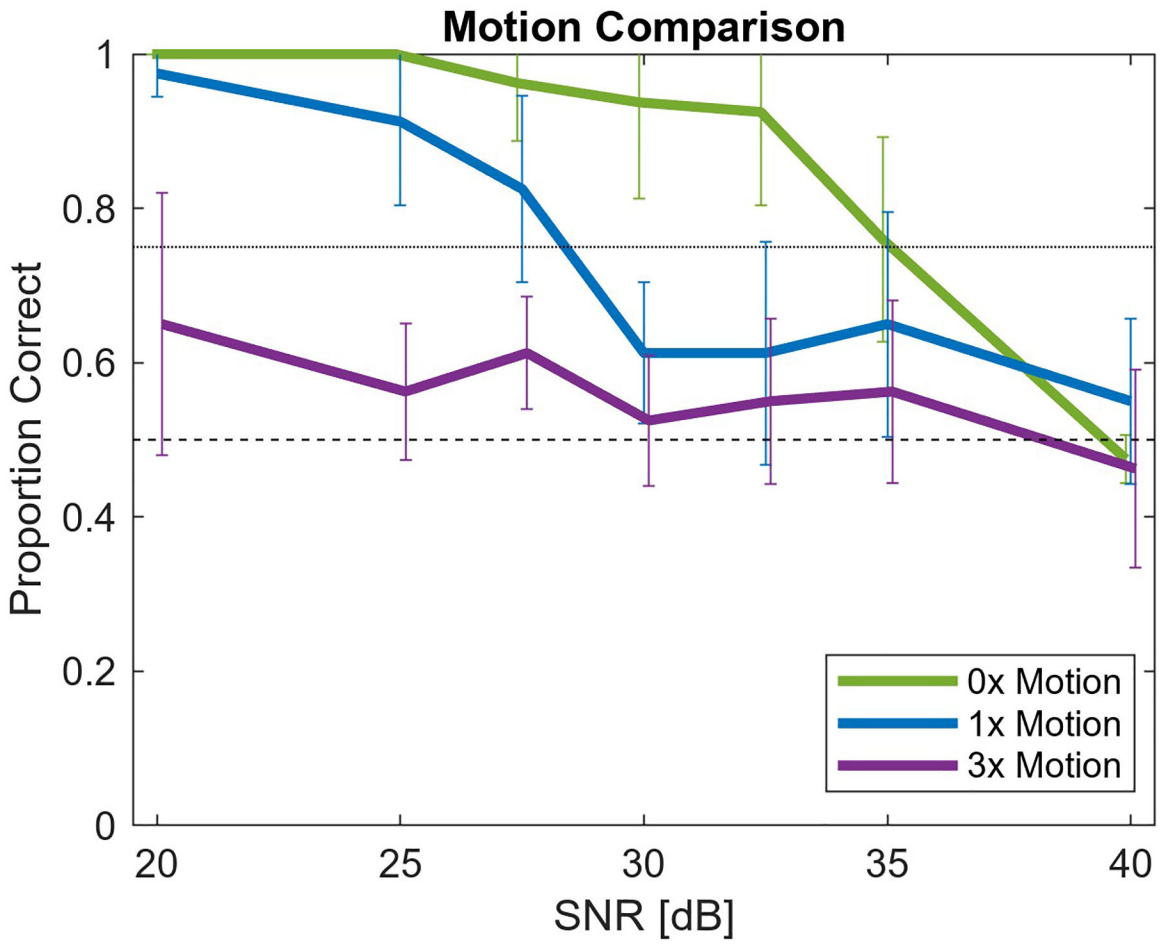

Fig. 4:

Proportion of correct identification of noise-added cine clips at 2 MHz as a function of temporal SNR. Three motion levels (stationary, standard, and exaggerated) are represented by the 0x, 1x, and 3x legend. With higher motion, the noise is less perceptible.

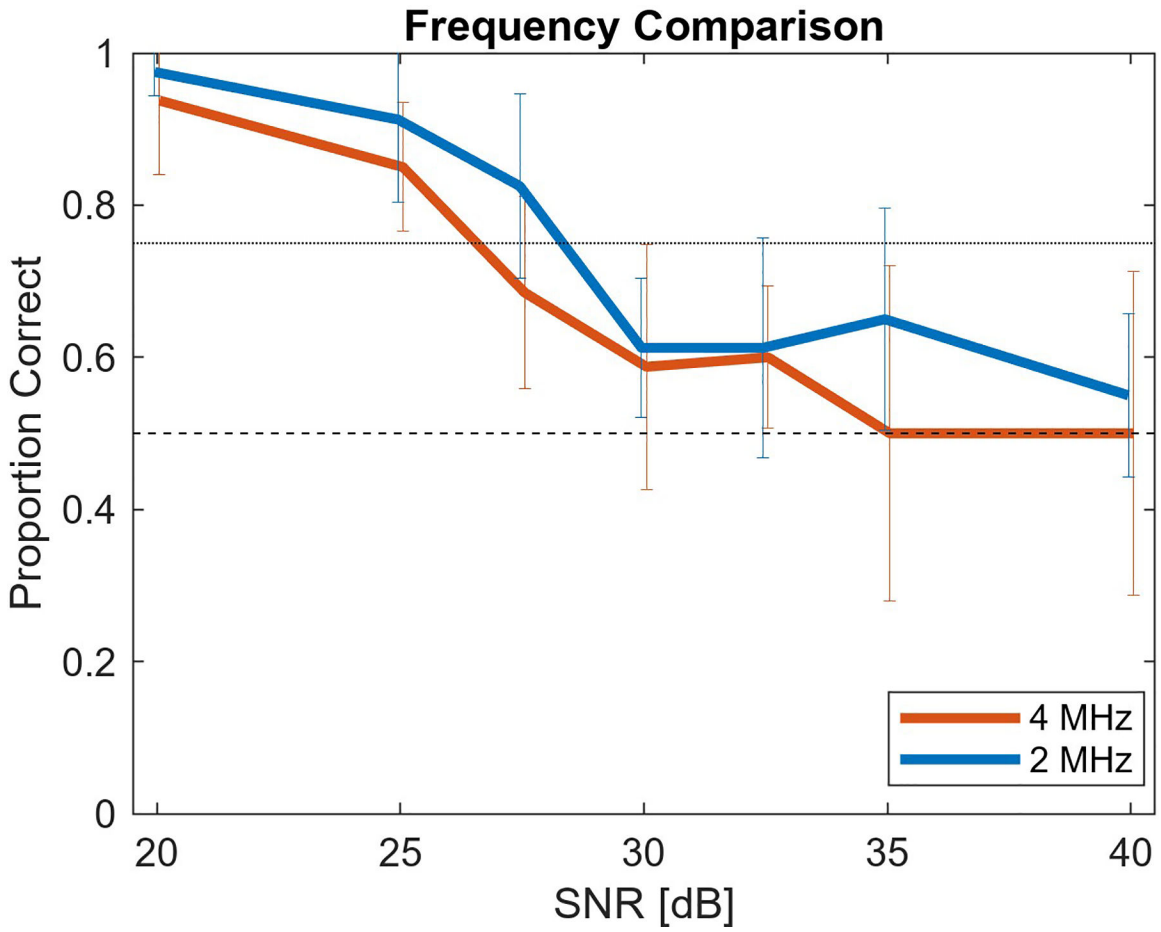

Fig. 5:
Proportion of correct identification of noise-added 1x motion level cine clips at two frequency levels, 2 MHz and 4 MHz. The frequency difference has a minimal impact on noise perception level.

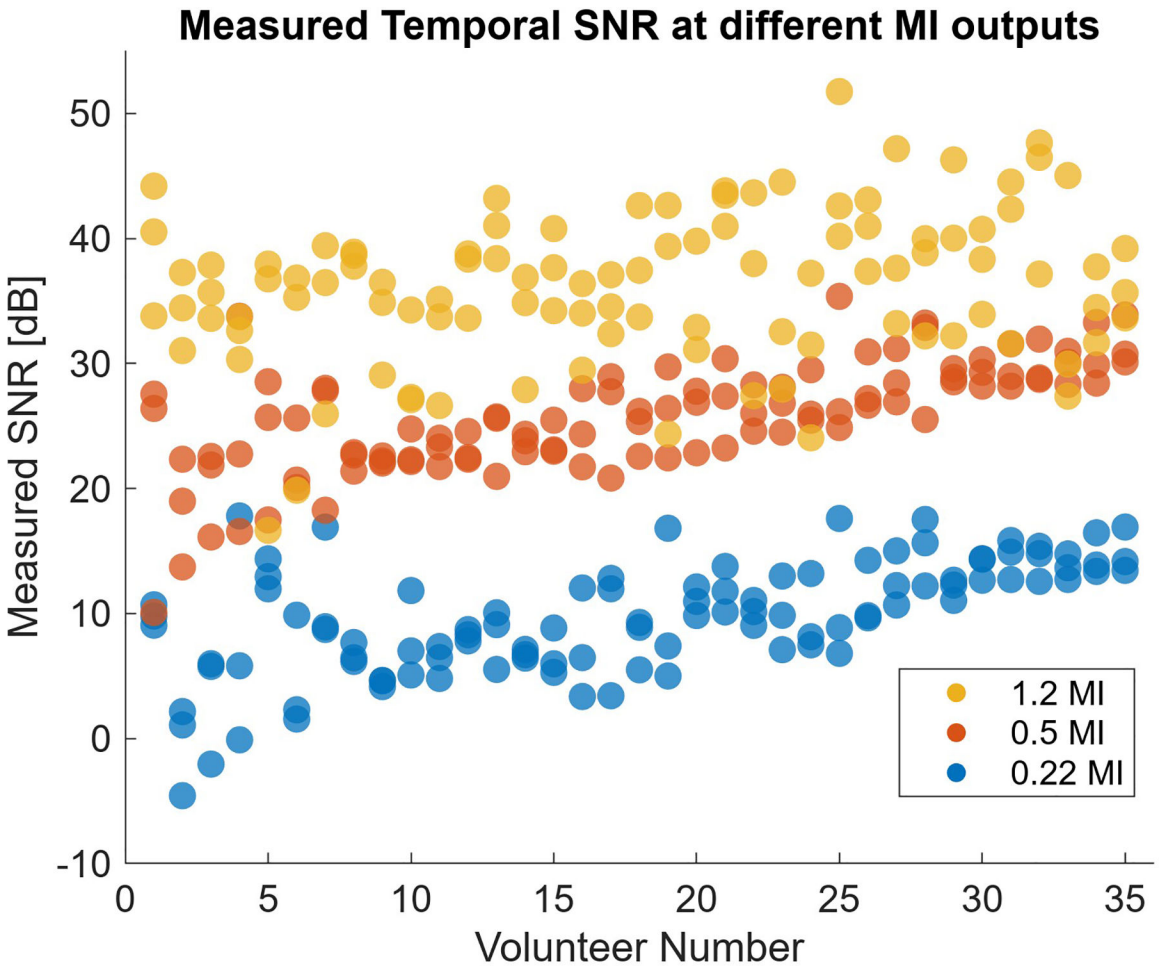

Fig. 6:
Observed SNR at different MI output levels in clinical output sweep. Higher output is generally associated with increased temporal SNR.

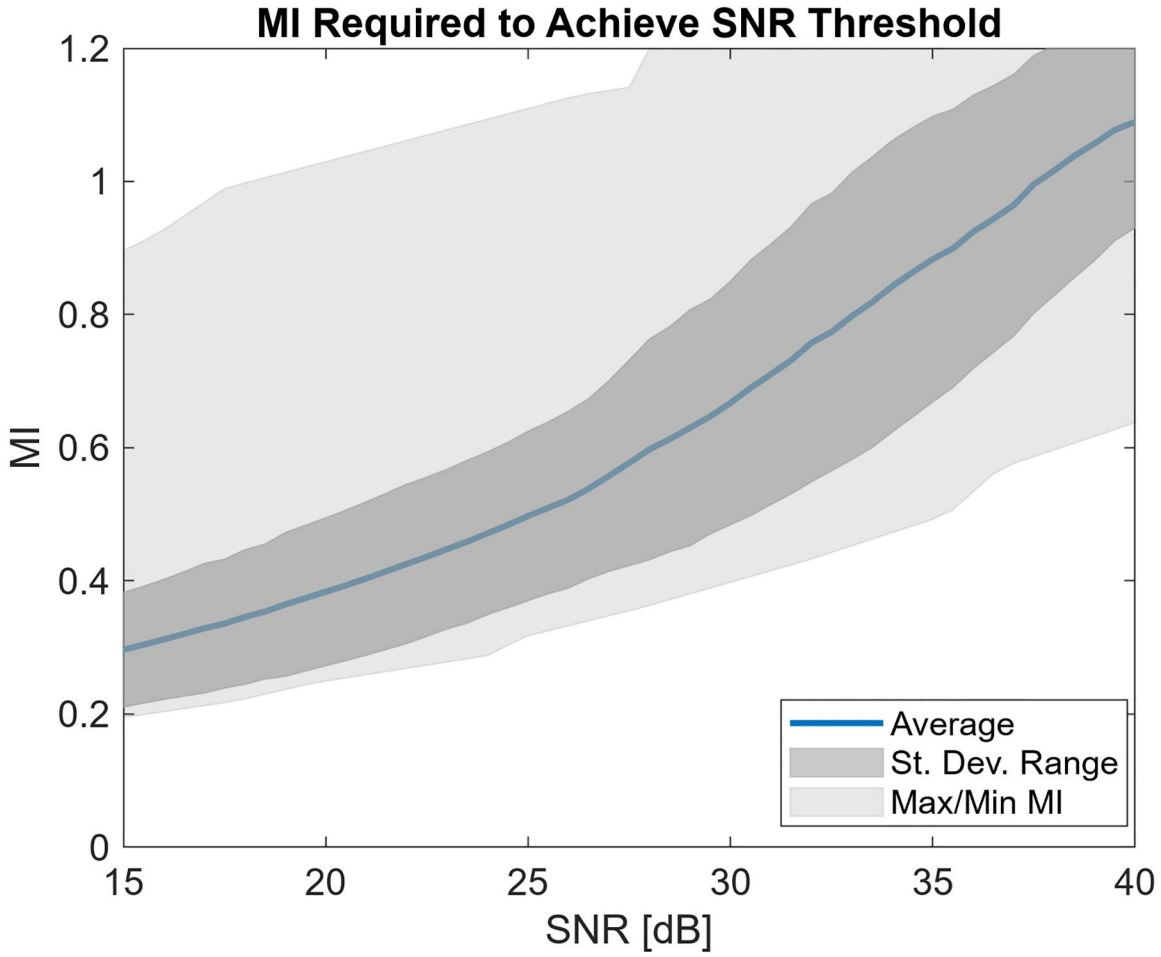

Fig. 7:

Estimated MI required to reach given temporal SNR imaging conditions. Increasing SNR quality requirements leads to higher recommended acoustic output levels.

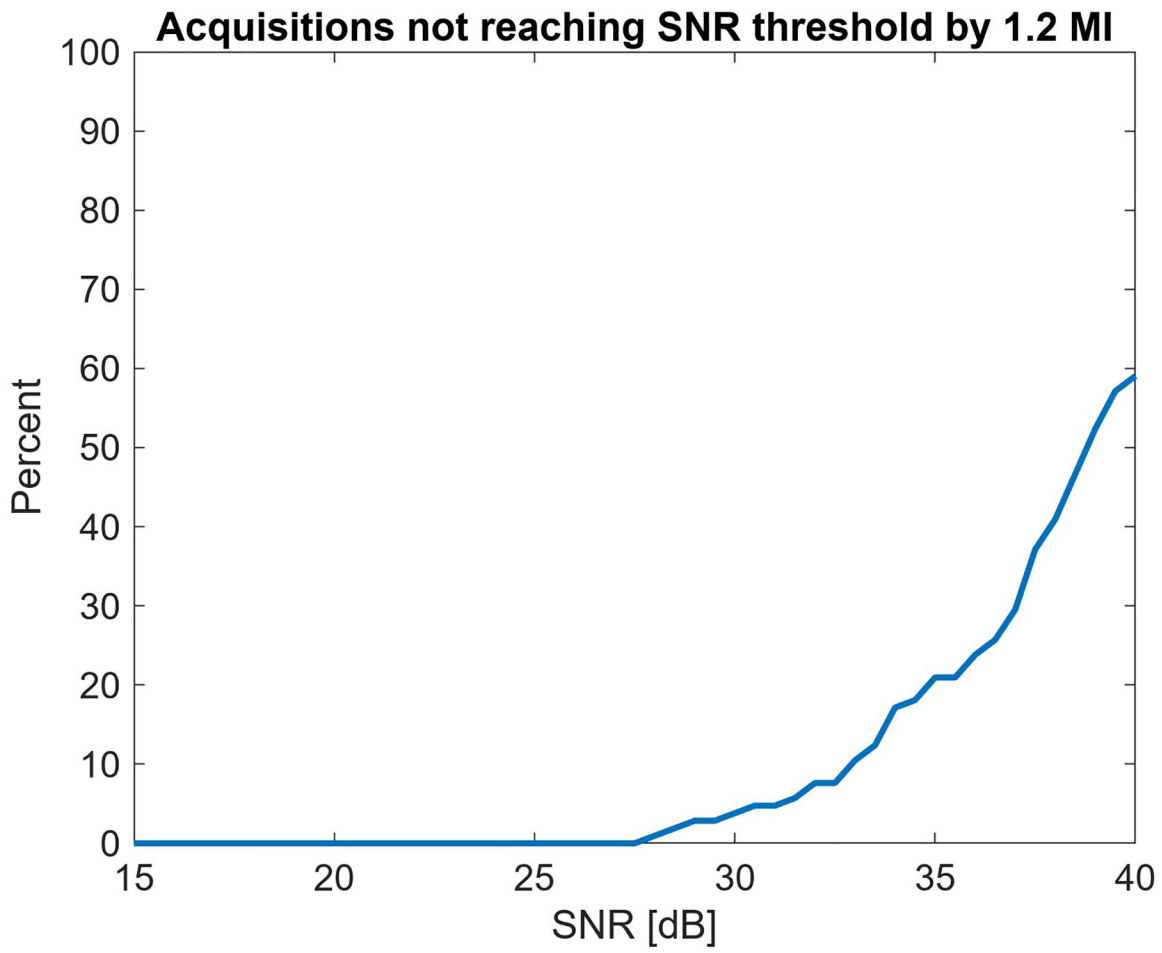

Fig. 8:

The maximum output of 1.2 MI in the collected data is not always sufficient for achieving a given SNR threshold. The first instances where this is observed are at 27 dB SNR, and by 39 dB a majority of scans are acoustic output limited.

Table I:

Temporal SNR defining noise perceptibility limit

Condition	Approximate SNR Perception Cutoff	Potential SNR Cutoff Range
2 MHz, 0x Motion	35 dB	Less than 40 dB, but greater than 32.5 dB
2 MHz, 1x Motion	28 dB	Less than 30 dB, but greater than 25 dB
4 MHz, 1x Motion	27 dB	Less than 30 dB, but greater than 25 dB
2 MHz, 3x Motion	<20 dB	Less than 20 dB

Table II:

Acoustic output required to achieve temporal noise perception thresholds

SNR Condition	Mean MI +/− St. Dev.	MI Range
20 dB [2 MHz, 3x Motion]	0.38 +/− 0.11	0.25 – 1.0
27 dB [4 MHz, 1x Motion]	0.56 +/− 0.14	0.34 – 1.1
28 dB [2 MHz, 1x Motion]	0.60 +/− 0.17	0.36 – 1.2
35 dB [2 MHz, 0x Motion]	0.88 +/− 0.21	0.49 – 1.2