Temporal Bone Dissection Simulation

D. Research Design and Methods
D.1 Overall Research Design

In section A, we established the justification and specific aims of this work. In section B we delineated the history preceding this work and the significance to otological training and healthcare. We established that only through multi-institutional validation studies can efficacy within the otological curriculum be determined. The preliminary studies section indicates that the system in its current state is already useful in the training of novice otologic surgeons. The overwhelming response by those at other leading institutions that have seen and interacted with the system lends support to this claim (see appendix). Below, we present our research methods to achieve the goals and objectives set forth in Section A.

This work will be accomplished through iteration of the following key areas:

- Structural Acquisition (CT, MR, microCT, High-Field MRI at the local level)
- Integration (data, system operations, file collection, and realism)
- Dissemination, (Remote site software, training, and physical system distribution),and
- Validation (user trials at local and remote sites)

All four areas will be pursued locally at OSU/OSC in Columbus. Remote sites are expected to assist only in defining demographics data of participating residents and learning standards, and to execute validation studies at their site. Under the direction of the PI, Dr. Wiet, the Co-Pi, Mr. Stredney will direct the organizational effort for the project. This includes overseeing all the key areas of data acquisition, integration, dissemination, and validation**.** The bulk of the work will proceed at the **local** level in Columbus. The remote institutions participating in the study will facilitate overall development of the standards for data to represent the anatomy (i.e., structures, pathologies, techniques) and direct validation studies at the remote site through identification of residents and by running validation studies. **NOTE: All participation by remote sites is purely voluntary. There is no compensation provided by this proposal.** Mr. Stredney will work directly with these individuals to facilitate the dissemination and evaluation of the system.

D.1.1 Domain (Wiet, Welling, Stredney)
Each site has established a domain expert who will develop criteria for structural data and pedagogical standards. Additional members from the Academies and NIDCD will be recruited. Collectively, we will deliberate and select common structural criteria and issues related to the process of technique to be integrated into the system. Participants will establish the requisite demograhic information for the studies. These individuals will assure that the system meets the general needs of their curriculum before validation studies are conducted at their site.

D.1.2 Acquisition (Wiet, Stredney, Murakami (CT), Powell (□CT), Schmalbrock (MR- High Field MR), (GRA-TBN) This group will oversee data acquisition standards and optimization for reconstruction.
A wide variability in data acquisition capabilities is expected from the various remote sites. The disseminated system has multiple data sets to perform dissection simulation that were acquired and integrated by the local team at OSU/OSC. The local team in Columbus expects to obtain 1 to 2 new data sets (bilateral)per year). In this way, we offer not only anatomical variance in the simulator, but also the possibility of participation for a remote site **without** requiring them to acquire and develop their own structural data. However, several sites have expressed interest in developing and integrating data with the understanding that these data will meet the general criteria for use in the simulator. We will provide acquisition parameters and protocols optimized for contrast enhancement and reconstruction, i.e., as close to isotropic voxels as possible. As stated above, we will provide sites with software to assist in the segmentation of their structural data. All developments will reflect emerging standards in data modeling that are being developed by other agencies, e.g. National Library of Medicine, and NIDCD's Temporal Bone Registry.

D.1.3 Integration (Wiet, Stredney, Shen, Bryan, Sessanna, Clymer, Krishnamurthy (Registration/Segmentation) GRA-TBN) Oversee all data integration and software development.

Similar to acquisition, we expect a wide range of technical capabilities at each remote site. Some sites are early adopters and have technologists active in simulation development. Other sites do not have an active group pursuing simulation development. Advanced sites may require only a simple download of executable software to run the simulator. Others sites will require assistance in planning the purchasing of their own system to participate. In an attempt to include as many sites as possible, we are requesting $60,000 for the purchase of three systems, one in year 3, 4 and 5. These will be integrated locally at OSU/OSC and will be shipped and installed at remote sites who wish to participate in the study, but cannot purchase their own equipment. It must be noted that we have based this price ($20,000) for each system to include all computing and interface equipment, as well as all shipping and handling charges. It is quite likely that some costs will vary in the 3-5 years hence. We have established close working relationships with various vendors to assure the purchase of a cost effective system design. During Phase III (see below) remote sites will be provided the equipment on loan for a fixed period of time that maximizes the number of institutions requesting participation.

D.1.4 Dissemination (Wiet, Stredney, Sessanna, Bryan, Weisenberger)

To expedite the transfer of information, we will establish a virtual collaboration between the local group and the remote participants. Information will be provided through a web site to appraise remote participants of updates and to request information (we currently maintain a web site to sample information from the remote sites regarding their interest, contact information, programs, pedagogical approaches, and computing environments). Periodic checks with remote sites will be made by phone. In addition, we will utilize national meetings such as the AAO-HNSF, to provide “Birds of a Feather” sessions, to train and communicate updates, as well as to seek consensus on developmental approaches. Due to the wide diversity in learning otological surgical techniques, we seek to create a forum for the exchange of ideas, and the establishment of approaches that serve the common need. Only through the open exchange, deliberation, integration, and evaluation by multiple sites, will the system be able to reflect the common approaches to otologic surgery training. Afterwards, if at all possible, idiosyncratic approaches will be integrated to provide further variability to the system.

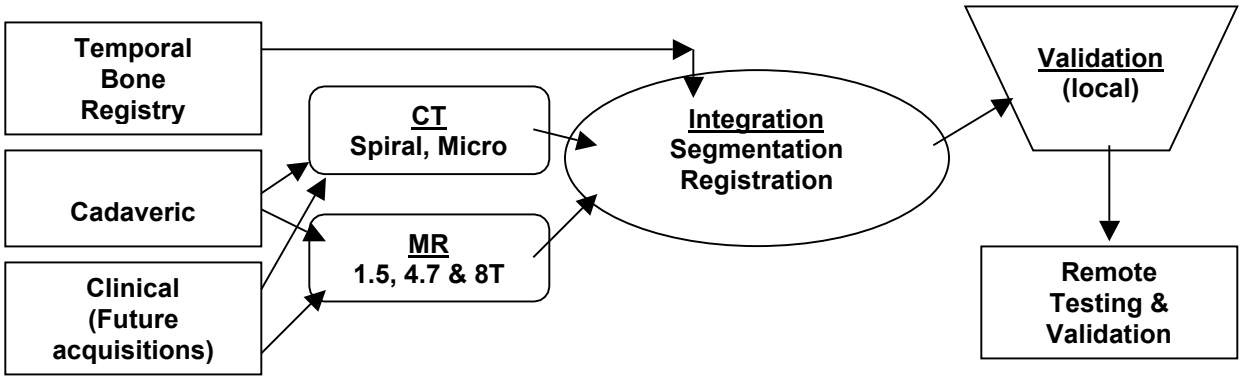
Remote sites that have expressed interest in the system and in participating in a multi-institutional study are diverse in their program size, pedagogical approaches, and technological capabilities. Some sites, who are developers and early adopters, are seeking to obtain the software immediately. They possess the technical capability to accept the system and run it on their existing equipment. Some sites currently possess little or no simulation capabilities. Our approach will be to engage all sites in the development of data standardization. Early adopters, who possess the technical requirements, will receive executable code within the first year. For those currently lacking the technical requirements, we will consult and recommend systems or subsets to complete a system for them to purchase, in Phase II (see Timetable) or wait until Phase III to obtain a system.

D.1.5 Validation (Wiet, Welling, Stredney, Weisenberger, Fernandez and Remote Participants)

We have established common criteria in data and testing standards for the study. We have refined our study design for our remote validation studies (see Section D). We will provide multiple mechanisms for data retrieval through direct network to the machine, or, in the event that no network connections are available, the software will provide capabilities to write the obtained data to external media, e.g., CD-RW.

D.2 TimeTable

The work will proceed in three phases. During the first year, **Phase I**, we will focus on migrating the current system to a more cost-effective portable platform. Working in collaboration with the remote sites and the academies, we will establish the prerequisite criteria to certify data sets for use in the simulator. In **Phase II**, or year 2, we will test the system under a local study. . In **Phase III**, or years 3- 5, we will disseminate the system to the remote sites participating in the study. We will train local representatives in the use of the system. Through all three phases, we will continue with the acquisition and integration of multi-resolution and multi-modal data to support more complex surgical procedures. We will explore the integration of data from the NIDCD temporal bone registry (See Nadol letter in appendix), and make recommendations for future processing and integration of registry data. In this way, we will maximize the access and use of the extensive data repositories in the temporal bone registry.



Data Flow Chart: Flow from acquisition sources, modalities, system integration, dissemination and validation

D.3 Structural Acquisition

To clarify, we define **structural acquisition** as the representation of the regional anatomy of the temporal bone acquired by traditional medical imaging techniques, i.e., computed tomography, magnetic resonance, macroscopic section, any discrete sampling that can be reconstructed. **Data collection** refers to demographic data of residents and experts using the system, as well as performance metrics that are captured during the system validation studies. The methods to collect these data are discussed under the section D.7 of validation**.**

The structural data that represents the regional anatomy for which the simulation is based must be of a resolution that provides precise and accurate structural qualities. To achieve this objective, we have selected unique and emerging imaging technologies, including custom mircroCT and high filed magnetic resonance imaging, unique to Ohio. We present below detailed descriptions of our imaging methodologies.

D. 4 Structural Acquisition Methods

The quality of the simulator is directly dependent on the quality of the structural representation used to model the visual, haptic, and aural modalities. We will continue to work with specialists developing emerging imaging technologies to acquire the highest quality of data for use in the simulator. In addition, we will provide recommendations for acquisition protocols optimized for reconstruction to remote sites that wish to develop local data sets that can be used within the simulator. We will obtain 1 – 2 fresh cadaveric specimens to acquire new multimodal, multi-scale data each year of the 5 year study. In this way we hope to have between 6 to 11 unique data sets (both left and right) by the end of the project to disseminate. We fully expect that additional sites pursuing simulations will wish to contribute to this number.

D.5 Local Acquisition of Structural Data

We will obtain multi-resolution structural data to support the precision required to simulate otologic procedures. Our multi-modal acquisitions include both magnetic resonance for soft tissue detail, and computed tomography for boney detail. We present our current methods here and discuss our plans to further integrate these data into the simulator. As mentioned above, we will investigate the integration of temporal bone registry data.

To obtain multimodal data sets, we will acquire data from fresh cadavers obtained through the Body Donation Program of the Department of Biomedical Informatics at The Ohio State University (OSU 2002). Insitu imaging will be performed to preserve the relationships with soft and boney tissue, and to characterize the orientation of the inner ear anatomy with respect to external anatomic structures. The temporal bones will then be excised bilaterally and imaged again with micro imaging techniques. Insitu Magnetic Resonance imaging (1.5Tesla) will be conducted at the Magnetic Resonance Facility at OSU, and Computed Tomography (CT) is conducted at Children's Hospital, Columbus. MR micro imaging data of the excised specimen will be acquired using the new 4.7T/40cm systems and the high-field 8T/80cm at OSU, and microCT will be conducted by the Biomedical Engineering Center at the Cleveland Clinic Foundation.

Insitu CT scans were acquired with a multi-detector (8 slice) CT scanner (Light Speed Ultra; GE Medical Systems, Milwaukee, Wis.). Specimens are scanned and reconstructed with the following parameters: section

thickness of 0.625 mm using an axial acquisition (scanner currently cannot do spiral acquisition with this section thickness), gantry rotation time of 1 second, x-ray tube voltage of 120 kV, x-ray tube current of 200 mA, and an imaging field of view of 10 cm. The imaging protocol yields 0.625 mm thick axial images of the temporal bone with a 512 x 512 image matrix. Volume resolution will therefore be 0.19 x 0.19 x 0.625 mm**.** In our initial study, in situ Magnetic Resonance Images (1.5Tesla) were acquired at the OSU Facility in about 20 minutes with a T2-weighted 3D fast spin echo sequence resulting in 0.35x0.7x0.7mm voxels.

Subsequent to insitu imaging, bilateral excision of the temporal bone was performed and the specimens were trimmed to contain the middle ear laterally, and the otic capsule and internal acoustic canal medially. Micro-computed tomography images of the excised temporal bone was obtained using a dedicated microradiography system consisting of a microfocal x-ray source (Phoenix X/ray, Wunstorf Germany) a 7-axis micro-positioning system and a high resolution X-ray image intensifier coupled to a scientific CCD camera (Thomson Tubes Electroniques, Cedex France) (See Figure 8). In order to safely expose the specimens to air, a 70% ethanol solution is used to exchange the formalin initially used to fix the specimens. One hundred and eighty 512 x 512 12-bit projection radiographs are collected in a circle at 1degree intervals around the specimen. The projection radiographs are acquired at 40 kV and 200 microA with the image intensiifier operating in 7-inch mode. The images are background corrected; the gray level intensities are scaled relative to that of air. A 512 x 512 x 512 volume is reconstructed using a modified Feldkamp con cone-beam reconstruction algorithm (Grass 2000). For the proposed research, in situ CT and specimen microCT images will be acquired in a similar fashion as in our initial studies.

The MRI data for the excised temporal bone were acquired on a GE 1.5 Tesla employing 3D fast-spin echo sequence with a TR of 2000ms TE 80ms using custom 1 inch RF receive coil. The data were reconstructed to an isotropic voxel resolution of 0.3x0.3x0.3mm. Initial high-field 8T images were acquired a small single strut TEM RF coil. Ultra-high resolution images were obtained in 2hrs with a 3D spin echo sequence using TR 300ms, TE16.6ms FOV 5.5x5.5cm and a 5cm slab, matrix 256x256x128 (voxel resolution 214x214x390μm, see Figure 9). In addition, gradient echo images were acquired in 20 minutes using TR=50ms, TE=5.3ms and about 90 degree flip angle. While allowing for significantly shorter scan time, gradient echo images were somewhat degraded by susceptibility artifacts.

In MR imaging, spatial resolution (i.e. the volume of each 3 dimensional image element or voxel) is limited predominantly by the available signal-to-noise ratio (SNR), but also by time constraints imposed by the spatial encoding process. In theory, SNR improves approximately quadratic with magnetic field strengths, however other confounding factors (RF inhomogeneity, susceptibility effects, longer T1 and shorter T2 and T2* relaxation) result in practical 5-10fold SNR increase between 1.5 and 8 Tesla, and 8T MRI has demonstrated potential for superior image quality (Bourekas 1999, Christoforidis 1999, Burgess 1999, Schmalbrock 2002, Abduljalil 2002, Christoforidis 2002). Furthermore, substantial SNR increase is achieved when small receive antennas are placed in close proximity of the anatomic structures to be imaged(e.g the voxel volume could be reduced more than 5-fold, in 1.5T specimen studies compared to in situ studies). Finally, very strong magnetic field gradients are required for ultra-high resolution MRI. The 8T human MRI system does not have such gradients thus restricting spatial resolution. However, such gradients are typically available with animal MRI systems.

For the research proposed in this application, a limited number of in situ studies will be acquired at 1.5T using optimized acquisition protocols that were previously developed at OSU (Kurucay 97, Stone 98, Schmalbrock 99, Massick 00, Schmalbrock 00). Insitu temporal bone imaging at 8T is severely hampered by RF inhomogeneity and susceptibility artifacts near air/bone/tissue interfaces are exaggerated at higher fields. Methods for overcoming these problems are a major research focus of the OSU 8T MRI research group, and 8T insitu imaging may be utilized for this work as novel techniques become available.

Excised specimens will be imaged at 4.7T and/or 8T. We were recently awarded funding for a 4.7T, 40 cm bore animal MRI system. Installation should be completed by December 2003. This system will be equipped with small volume RF coils and micro-imaging gradients optimally suited for studies of temporal bone

specimen. Susceptibility artifacts and RF inhomogeneity problems encountered at 8T will be insignificant or at least less severe at 4.7T. We have previously identified motion compensated steady state gradient echo

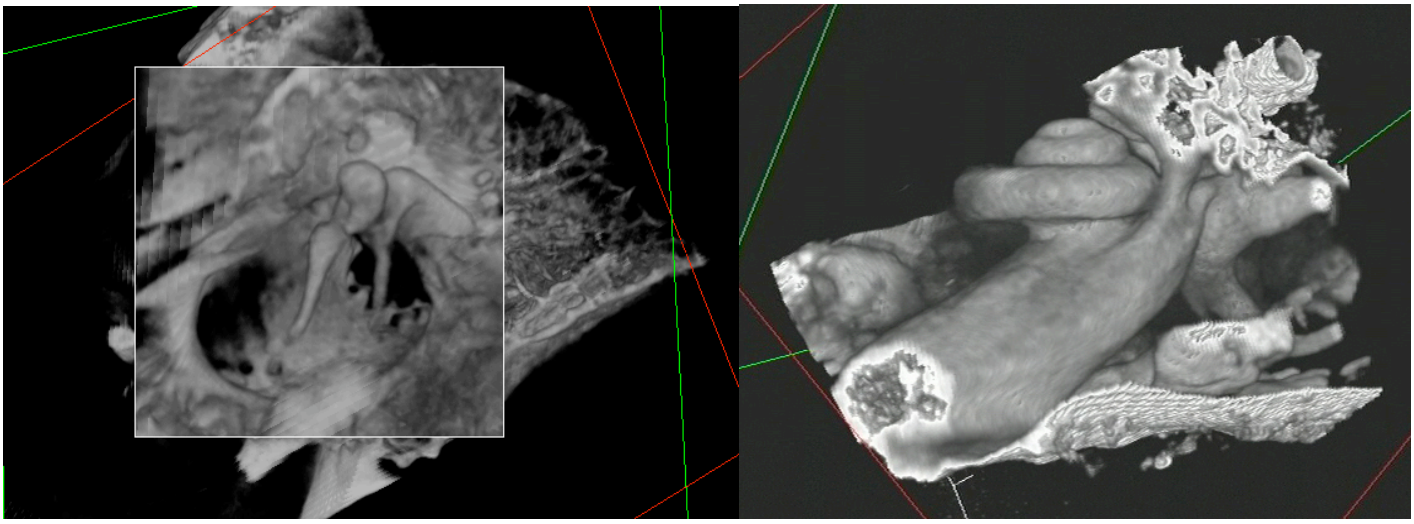


Figure 8: CT Left: Relationships of ossicles are highlighted. Right: Demonstration of vestibulocochlear nerve.

(Kurucay 97) or fast-recovery gradient echo sequences (Schmalbrock 00) to produce optimal results for temporal bone imaging. These sequences are currently not available with the BRUKER console on the 4.7T system and will be implemented for these studies. We expect to achieve isotropic resolution between 100-200 m^3 with this setup. Because 8T MRI can potentially achieve yet higher spatial resolution than 4.7T, and because the 8T whole body system can be used for human in vivo studies in the future, we will also continue to improve acquisition methods for imaging temporal bone specimen. We have already implemented a better 3D gradient echo technique. Since both the 4.7T and 8T MRI units are controlled by BRUKER consoles, the implementation of motion compensated steady state gradient echo (Kurucay 97) or fast-recovery gradient echo sequences (Schmalbrock 00) at 4.7T will be available at 8T as well. Finally, several coil designs for small specimen studies with improved SNR at 8T are already underway, and we have demonstrated that SNR can be improved and scan times reduced at 8T by using multi-channel receive, multi-port TEM coils with SENSE and SMASH parallel imaging (Abduljalil 03, Schmalbrock 03).

Finally, since combination of multi-resolution, multi-modality data is important to this project, it is essential to validate the image geometry between 8T, 4.7T and 1.5T MRI as well as CT and microCT. While clinical MRI and CT instruments routinely employ appropriate calibration and geometry distortion correction, the precision of the correction has to be confirmed for small specimens. The 4.7T and 8T MRI systems use only simple linear calibration. Correction to account for non-linear deformation due to gradient non-linearity has not been evaluated but is evident in image examples. We will quantitatively evaluate these issues e.g. by using phantoms with regularly spaced grids (see Fig 11) and implement appropriate correction during image acquisition and reconstruction. To validate the integrity of the image geometry, we will also develop a fiduciary/grid system to be integrated in the sample holder for multi-modality, multi resolution studies.

D.6 Integration of Structural Data

To register the multiple resolution and multimodality images, the approach shown in Fig. 10 will be used. The lower resolution images from the commercially available CT and 1.5T MRI whole body scanners will be used as rigid body references for the higher resolution 4.7T or 8T MRI images and CT images. Mutual information techniques, which are known to perform well in multimodality applications (Viola 1995, Wells 1996,Maes 1997, Pluim 2000, Pluim 2003), will be used for 3-D rigid body registration between the in situ images from 1.5T MRI and the whole body CT scanners. Because the 4.7T MRI, 8T MRI and CT imagers are custom-built in-house from components, non-rigid registration may be required to morph high resolution images from these scanners to the lower resolution landmarks identified in the image data from the 1.5T MRI and the whole body CT scanners. This post-processing technique will correct artifacts from ultra high-field susceptibility effects and

gradient nonlinearities in the custom built MRI scanners and beam hardening effects in the custom built CT scanner

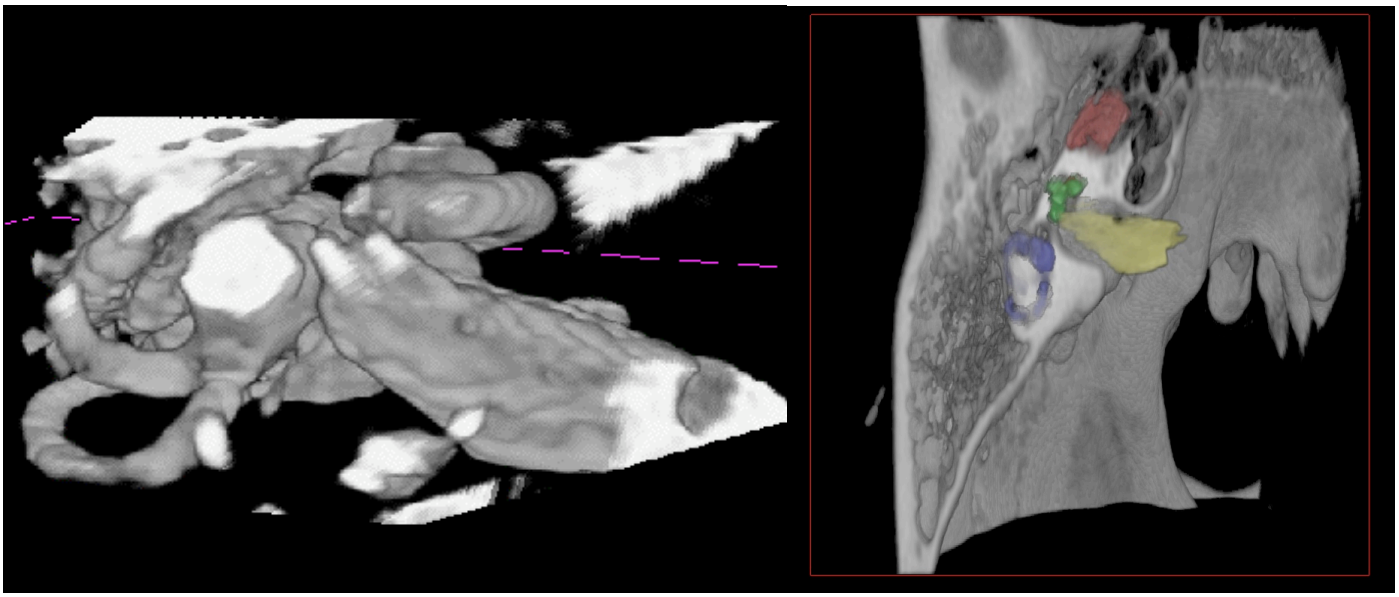


Figure 9:Left = High-Filed MRI of vestibulocochlear nerve (Cochlea= upper-center, semicircular canals = lower left (secioned) Right = Multimodal/multiscale data merged (spiral CT with MRI from High Field 8T(color)) NOTE: retraction of CNVIII due to excision

.

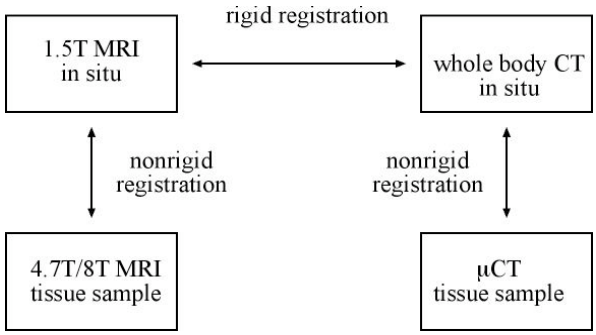


Figure 10:Image registration strategy

D.6.1Validation of Registration

We have commenced validation of our registration by using a geometrically regular phantom (See Fig. 11), a section of a DNA micro-array with wells of 20 microliters in volume. The irregular cuts on the edges provide unique features for registration. For our cadaver studies, we will employ implanted landmarks in the insitu and excised tissue. The distortion evident in the images will require a straightforward morphing algorithm. This transform will be smooth (slowly varying) because the distortion is primarily due to a relatively gradual shift in the position of the 8T MRI phantom image gradient field nonlinearities. The phantom is currently being imaged in the whole boy (1.5 MRI and clinical CT) scanners as well as the 4.7T MRI. Image acquisition parameters are such that image resolution and contrast for each scanner and modality will be comparable to those typical for the cadaver and tissue images, using contrast agent as appropriate. The phantom will be used to verify rigid registration between the low resolution 1.5T and whole body CT images, using the measured spacing of the centroids (in three dimensions) of the well array and the root mean square (RMS) displacement between the centroids in the two imaging modalities as a figure of merit.

To investigate the need for non-rigid registration of the high resolution images of the well phantom from the custom scanner to the lower resolution references, images from the 4.7T MRI and 8T MRI will be registered as rigid bodies to the corresponding images from the standard imagers. RMS displacement between matched centroid will be used as a measure of non-rigid distortion of the high-resolution images.

If nonrigid registration is deemed necessary, coarsely distributed wells in the array will be used as landmarks for the image morphing algorithm. The RMS displacement of the centroid of the wells after non-rigid registration will be used as a figure of merit. Experiments varying the concentration of landmark position and the ability of the morphing algorithm to provide acceptable registration will be studied. Results will be used to determine the concentration of landmarks required for nonrigid registration of the in situ and excised tissue images.

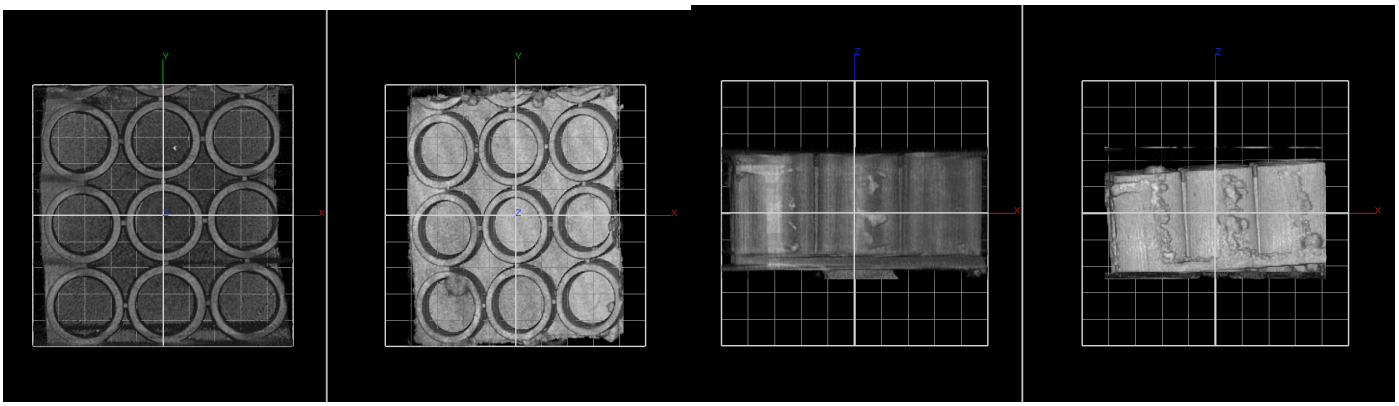


Figure 11: Geometrically regular phantom. Left = CT/8TMRI . Right = CT/8T MR

As verification of registration of insitu and excised tissue images, markers visible in both MRI and CT will be surgically embedded in the cadaver before scanning.For example, we are considering .5 inch nylon threaded tube fittings filled with fluid and capped at the ends. These would be visible by the CT, and both the fluid center and impression left upon the soft tissue of the external auditory canal would be visible in the MR,. Markers will be placed in locations so that they remain embedded in the excised tissue for high resolution scans. For low-resolution MRI and CT registration, markers will be cropped or deleted from the image set, but location within the image volume will be noted. Image registration of the volumes (with markers absent) will be registered. After registration is complete, the relative positions of the markers will be calculated. Figure of merit for the insitu and tissue image registration will be RMS displacement between the centroids of the calculated marker positions.

For finer registration verification of the cadaver and excised tissue images, experts will identify landmarks that are visible in all sets of images. These landmarks will not be used during the registration process, but their locations after registration will be noted in each set of images. Figure of merit will be the RMS displacement between identified landmarks in the registered image sets.

D.6.2 Image segmentation:
Due to anatomical variance and precisions of data acquisitions, there is no “gold standard” for segmentation. Others (Kaus 2001) have established the ‘gold standard’ as derived by experts employing manual segmentations using interactive techniques similar to those we presented above in our preliminary studies. We will define our standard as only those voxels that have been identified by 3 out of 4 experts as inclusive to the structure. We will employ a volume comparison of matched tissues in the different imaging modalities and resolution with data segmented and approved by experts. We will employ the low-resolution images (calculate macro volume of visible tissue regions in each) as the “gold standard” and compare to matched regions in the higher resolution images. We will compare the matched segmented volumes in the 1.5T and whole body CT image sets to each other, then each with a manual segmentation performed by an expert.

Through the exploration of evolving image processing API’s such as, VTK (VTK 2003), ITK (Yoo 2002. ITK 2003) and the use of our segmentation and registration software described above (see Figure 9), we will process and integrate the multimodal data sets for use in the simulation. In collaboration with our remote colleagues and through the exploitation of open source software, including our own, we will determine and post best practice methodologies to expedite future data integration.

D.6.3 Improve Visual and Dynamic Realism

To improve the visual realism (in conjunction with improved data acquisition), we will exploit the greater processing power of current and emerging commodity systems and graphics accelerator cards. The new line of high-end consumer class graphics cards offer unprecedented performance as well as highly programmable vertex and pixel pipelines. This allows us to move a portion of the computational problem to the GPU (Graphics Processing Unit), freeing the main CPU to perform other tasks. Several authors have recently presented on hardware assistedvolume rendering, including uploading the workload of compression, filtering, and segmentation onto the GPU (Kruger 2003, Schneider 2003, Hadwiger 2003, and Viola 2003). We will investigate packing gradient, opacity, and segmentation information into a single 4-channel volume and use the graphics hardware to render more realistic representations of the data. This includes more accurate diffuse (directional) shading, specular components for wet surfaces, and fluid effects (Westerman 1998, Dachille 1998, Van Gelder 1996).

We will carefully explore the introduction of secondary characteristics. **Our focus will remain real-time performance,** but increasing levels of realism are inevitable with emerging hardware. Initial efforts will include more realistic lighting, e.g., shadows, to more clearly depict form. Later in the project, we will investigate approximations of the underlying mechanics of fluid effects, bone dust (John 2001, Agus 2003), and soft-tissue deformations. This includes schematic representations of specular effects and environmental mapping for fluid representations, and simplified particle systems. For bone dust, we will explore the propagation of voxels along trajectories perpendicular to the axis of rotation and tangent to the drilling burr. We will only introduce these secondary characteristics if the schema provide real-time performance.

The simulation of realistic, complex structural dynamics in real-time is non-trivial. As we integrate soft tissue structures from the magnetic resonance data, it becomes necessary to investigate more accurate modeling of tissue dynamics. Soft tissue modeling will be designed to give appropriate visual and haptic cues when probed. The task of soft-tissue modeling in this region, as compared to say, laproscopic simulation, is somewhat easier for three reasons. First, most of the soft tissues in the region are critical structures and need to be avoided. Second, it is unlikely that any haptic feedback would be received when soft-tissue is accidentally contacted by a drilling burr. Second, these structures are often surrounded completely or partially in bone, thus their deformations are constrained. Therefore we will examine and implement methods that will provide an effective but efficient level of realism and maintain real-time performance.

Researchers have applied physically based methods to simulate realistic deformations. For example, Terzopoulos et al. (Terzopoulos 1987) used the Finite Element Method (FEM) to model elastically deformable objects. Keeve et al. (Keeve 1996) developed an anatomy based model to perform facial tissue deformation for surgical simulations. Chen et al (Chen 1992] developed a biomechanical FEM considering skeleton kinematics and physiological effect for muscle deformation. In the literature, two approaches are generally used to perform model deformations. One is based on mass-spring models (Christensen 1997, Promayon 1996, Provot 1995) and the other is based on FEM simulations (Nielsen 1996, Lee 1995, Metaxas 1992). The mass-spring approach is simple and fast, but offers less visual realism due to its C0 continuous nature. The finite element methods can provide C1 or C2 continuous, and thus are more accurate. However, they are much slower to compute. More recent developments point to schema that provide real-time performance using volumetric data(Park 2002).

To simulate the soft tissues in our temporal bone surgical simulations, new tissue deformation algorithms that can achieve a balance between visual realism and system interactivity need to be developed. To achieve this goal, our work will focus on three parts: (1) interactive generation of volumetric mesh (2) hybrid linear and high order finite element computation model (3) interactive rendering of FEM volume results. The semi-automatic 3D image segmentation algorithm mentioned above will be used to identify soft tissues from the high-resolution imagery data to create an initial mesh. At run time, this mesh will be modified by the operations. Soft tissues can be cut away, or be deformed. Non-linear property and topology on the deformation of soft tissues will be simulated using the FEM method. Since non-linear properties often appear in the local regions around the operation position, high order FEM will be applied only to localized regions in the model with frequent dynamic updates, while other regions can remain constant. We will develop a computational model that can deal with

this adaptive computation in multi-scale regions to realize this idea. Since a majority of the computation will employ linear and topology constant FEM in the non-critical are, interactivity can be expected.

Using the programmable aspects of the latest commodity graphics hardware we will explore integration of advanced lighting effects such as realistic form shading with specular highlights. Recently there has been extensive research done in the development of advanced volumetric lighting techniques using programmable graphics hardware (Engel, 2001)
Secondary characteristics such as fluid effects can be represented in multiple ways. We will explore the use of a planer iso-surface to represent the pooling of water or blood. Additionally, programmable pixel and vertex shaders can be used to perturb specular effects to simulate water flowing over the volumetric surface. Advanced multi-texturing techniques can be used to overlay blood flow effects over the volumetric surface by cycling the multi-texture lookup coordinates. Lastly, we can use simplified particle systems to represent fluid droplets during splashings in open space. For bone dust, we will explore the propagation of voxels along trajectories perpendicular to the axis of rotation and tangent to the drilling burr. (John 2001, Agus 2003).

The simulation of realistic, complex structural dynamics in real-time is non-trivial. The task of soft-tissue modeling in this region, as compared to say, laproscopic simulation, is somewhat easier for three reasons. First, most of the soft tissues in the region are critical structures and need to be avoided. Second, it is unlikely that any haptic feedback would be received when soft-tissue is accidentally contacted by a drilling burr. Only visual feedback would be presented. Third, these structures are often surrounded completely or partially in bone, thus their deformations are highly constrained. Therefore we will examine and implement methods that will provide an effective but efficient level of realism and maintain real-time performance.

One method we will explore is to segment the soft tissue structures and represent these structures as an iso-surface. We will then merge iso-surface representation of these dynamic structures with the volumetric representation of the static bone structures. The vertices of the iso-surface can be used as elements in a spring-mass, ChainMail, or basic linear Finite Element Method (FEM) dynamic simulation. The constraints on each element can be mapped to the existence or non-existing of corresponding voxels that represent static bone material. The soft tissue deformation in this type of simulation is seen as only a “cue” to the user and does not in any way need to be physically accurate, just physically plausible. Additionally, the deformations are highly constrained because of the surrounding bone material. Another factor effecting performance is that most the rendering computation has been uploaded to the GPU allowing us to use most the CPU power for dynamic simulation and force feedback calculation. We will explore several methods such as those described in (Nienhuys, 2001) or tetrahedron based methods similar to those purposed by (Bielser, 2000) or [(Bielser, 1999).

D.6.4 Multi-resolution data

We are currently integrating data sets acquired from microCT and High-Field MRI (see Figures 9, 10) above) to support emulation of increasing or decreasing the magnification of the simulation. This requires the manipulation of the multi-resolution structural data through level–of-detail volume rendering. We are exploring several techniques that would allow for the efficient paging and correlation of volume data for this effort (LaMar 1999, Weiler 2000).

Several key issues need to be addressed when employing multi-resolution data sets in the simulation. First of all, to guarantee a smooth transition and blending of data at different resolutions when the user zooms in and out of a local region, thus emulating the change in magnification of the surgical scope. Avoiding abrupt changes of data resolution and eliminating boundary seams of different resolutions are necessary to ensure interactive performance. Previously, Weiler et. al., proposed to copy the voxels on the boundary of low resolution volumes to their high resolution counterparts to eliminate the seams. However, this technique requires the sizes of the high-resolution and low-resolution volumes to be maintained at a fixed ratio, which is not the case in our application. We will examine several approaches most appropriately suited to this application and select those that best assure real-time performance. Secondly, it is important to maintain data consistency. At any given instant during run-time, only data at a particular resolution will be modified by the user. However, to ensure that a correlation exists between data of different resolutions, simultaneous updates

of the data at all levels need to be accomplished in real-time. Overwriting data in either the texture memory or main memory can be both computationally expensive and adversely affect performance. Subsequently, we will investigate efficient data structures and indexing schemes to link between data of different levels of detail to address this issue. We will design asynchronous data read/write schemes to make the process transparent to the user. Finally, it is imperative to ensure a time-critical rendering of multi-resolution data. To balance between the visual realism and the run-time rendering speed, care must be taken in selecting the level of detail in different regions. If high resolution data are used globally to represent the model, real-time frame rates may not be achievable, and performance will be adversely affected. On the other hand, an overly aggressive use of low-resolution data can both adversely affect the accuracy of the model and the resulting image quality. A run-time algorithm is required to automatically select data of different resolutions to guarantee an interactive rendering speed while maximizing the model accuracy and rendering quality.

We will develop volume rendering algorithms that provide level of detail (LOD) for large scale multi-resolution spatial and temporal volume visualization. To estimate the rendering time for data of various spatial and temporal resolutions, robust performance models will be developed. We will also define the benefit functions to reflect the spatial and temporal data coherence and application-specific space-time importance metrics. To optimize the rendering performance and visualization quality, our algorithm will extend a texture hardware based predictive-reactive LOD selection algorithm that we developed (Li 2001, 2002). We do not make a global decision for the LODs in a single optimization run as found in (Funkhouser93). To improve the accuracy of performance prediction, we subdivide the volume into small blocks, and determine the LOD for each block using the performance statistics gathered at run time. We have demonstrated that we can control the rendering speed with less than 5% of error from the user's target frame rate while maximizing the rendering quality. Integration of this algorithm into the dissection simulation, as well as evaluations of the algorithm’s utility for maintaining real-time performance will be conducted.

To extend this algorithm for time-dependent data, several fundamental changes will be needed. Instead of subdividing the domain into three-dimensional blocks, we will treat the time-dependent data as a collection of four-dimensional (space+time) blocks in order to consider temporal coherence. The performance model should also take into account the effect of using different temporal resolutions in different time intervals to the overall animation speed. We will consider various acceleration schemes to optimize our performance. We will also survey and select schemes such as data I/O, network transmission delay, as well as wavelet reconstruction cost.

D.6.5 Time-Series Data Collection and Real-Time Playback

It is necessary to record the intermediate results during the surgical simulation so that the user’s performance can be reviewed and quantified with the performance of his/her peers as well as of the expert. Generally speaking, there are two types of recordable intermediate results. The first is the time-series volume data in which each time step represents the state of the 3D volume at a particular instant during the simulation. The second is the process of the surgery, such as the path of the surgical tools and the amount of bone or soft tissue being removed along the tool path. Storing both types of intermediate results requires efficient run-time algorithms and data compression schemes that can operate asynchronously with the other operators in the surgical simulation without affecting the system performance. In addition, effective data structures are required that can exploit the temporal coherence between the volumes in consecutive time steps so that real time playback at the post-processing stage becomes possible. Although spatial data encoding for visualization applications has been intensively studied, research on efficient data encoding schemes for three-dimensional volume data that evolves in time is still in its infancy. Previously, researchers have proposed various algorithms to encode three-dimensional time-varying data. Guthe et al [Guthe01] used a windowed motion compensation scheme to match the neighboring volume blocks in adjacent time steps and then employed wavelet transform to encode the differences. Lum et al [Lum01] used DCT to encode individual voxels along the time dimension, combined with color table lookup to animate the volumes. Sohn et al [Bajaj02] used wavelet transforms to create intra-coded volumes and applied difference encoding to compress the time-dependent volumes. While all these methods can achieve considerable compression rates, they require a pre-defined quality threshold to encode the data, and thus are lossy by nature. In addition, their schemes do not provide the functionality to enable interactive visualization of data defined at arbitrary spatial and temporal resolutions.

We will investigate alternative data encoding schemes to facilitate interactive browsing of time-series data sets. In addition to data compression, an important goal of our research is to de-correlate the underlying time-dependent data produced from the surgical simulations into a range of spatial and temporal levels of detail, so that the user can freely choose the desired resolutions for efficient playback and detailed analysis. Toward these goals, a lossless encoding scheme will be used. In this research, we will investigate the use of 4-D wavelet transform. Previously, researchers have proposed to use three-dimensional wavelet transform to perform video encoding and compression [Khalil99, Taubman94, Kim93]. Compared to the conventional MPEG approaches, wavelet transform of videos can produce less visual artifacts. Wavelet transforms also allow for hierarchical transmission of data, both in space and time. To apply wavelet transform for encoding time-varying scientific data, we can first perform a 3D wavelet transform to distill each volume at individual time steps into coefficients that correspond to different spatial scales. A one dimensional wavelet transform along the time dimension will follow to build a temporal hierarchy. The result of the transform will be a collection of coefficients corresponding to different spatial and temporal scales. These coefficients will be organized into an efficient indexing structure so that coefficients satisfying the user's quality criteria can be easily retrieved at run time.

To store the wavelet coefficients, since we intend to support interactive browsing of data at arbitrary resolutions, hierarchical data structures will be needed. Traditionally, image encoding schemes first partition the data into blocks, where transform encoding is applied to each block and the resulting coefficients corresponding to different scales are stored together. For image compression and progressive transmission applications, the organization of block coefficients is often not a major concern since the overall size of the transformed coefficients is typically small enough to keep in main memory to allow for random access. For large scale volume data sets, however, the size of resulting coefficients is often quite large, and thus it is imperative to store them in a manner that allows efficient retrieval. Previously, we have developed a data structure, called Time-Space Partitioning (TSP) tree, which allows for more efficient search of spatial and temporal coherence [Shen99]. In essence, the skeleton of a TSP tree is a standard complete octree, which recursively subdivides the volume spatially until all subvolumes reach a predefined minimum size. To store the temporal information, each TSP tree node itself is a binary tree, where each tree node represents a different time span for the same subvolume in the spatial domain.In this project, we will extend TSP trees to incorporate data encoding schemes facilitating progressive data transmission and interactive data browsing based on user-defined spatial and temporal quality threshold. The primarily function of the TSP tree is to perform efficient traversals of regions with different spatial and temporal coherence. As described above, we perform a wavelet transform in the spatial domain first, followed by a 1D wavelet transform in the time dimension for each data block. The spatial and temporal wavelet transform coefficients can be stored in the TSP tree. The TSP trees will be stored and managed in an out-of-core manner. At run time, based on the user desired quality or performance goals, wavelet coefficients both in space and time will be selectively transmitted, which can be used for downstream visualization processing.

D. 7 Validation

D.7.1 Study Design

We have designed this study mirroring the training and evaluation processes that have been in place at our institution for the past ten years. The studies will be conducted under the IRB approved protocol, “The Investigation of the Use and Efficacy of Advanced and Virtual Computer Interfaces for Biomedical Applications, 94H022, approved, August 11, 2003) (see Section E) Data will be collected using a combination of metrics stored to computer file during simulation and a written instrument/diaries to evaluate the system and its utility (see appendix). We will work with our remote collaborators to either adopt or adapt this protocol at their institution.

Only three individuals enter the residency program per year at The Ohio State University's Department of Otolaryngology. These three residents enter a five-year study program. During the first year, the residents enter a general surgical study and technically are not designated as Ear, Nose and Throat (ENT) residents. At most training programs across the country, residents attend a lecture series and temporal bone dissection course early on in their training. At our institution, this occurs during the late Spring/early Fall of post-graduate year 3 (PGY3). Our residents undergo a formal temporal bone dissection course that primarily involves lecture and actual dissection of human cadaveric temporal bones using the Temporal Bone Surgical Dissection

Manual (Nelson 1991). An evaluation is performed at the conclusion of the course using an objective, quantifiable grading of a dissected cadaveric temporal bone. The instrument used for this evaluation is the Welling Exam (See Welling Exam in Appendix). The Welling scale has a minimum of 15 points and a maximum of 36 points. The variability of previous scores using this method has ranged from 60 to 90 percent. This objective and quantifiable examination tool will serve as our pretest and posttest measurements to assess subject performance.

After receiving their basic training in temporal bone dissection, residents in the following three years, PGY3-5, are still required to attend the lectures, but are not required to attend the temporal bone dissection labs. Performance levels may differ. This difference may be expressed in the abilities of each resident and may be attributed to the number of exposures to temporal bone surgeries, including witnessing, assisting, and performing as well as self directed, un-mentored additional time in the temporal bone dissection laboratory practicing on cadaveric specimens

For purposes of this study, we therefore propose implementing a multi institutional basic exploratory pretest-posttest comparison of improvement in temporal bone dissection performance. This experimental design will measure a percent change in temporal bone score regarding the use of the simulator as a supplemental learning tool compared to additional time in the temporal bone laboratory. This design will take the following form:

PGY3	PGY4	PGY5
R 0 X1 0	R 0 X1 0	R 0X1 0
R 0 X2 0	R 0 X2 0	R 0 X2 0
R 0 X3 0	R 0 X3 0	R 0 X3 0

R = random selection
0 = pretest/posttest (Note: the pretest and posttest are exactly the same examination.)
X1 = (Control Group) No supplemental sessions
X2 = 6hrs of supplemental un-mentored traditional temporal bone lab sessions
X3 = 6hrs of supplemental un-mentored simulator temporal bone sessions
The independent variable of interest will be the type of supplemental training session (X1, X2, or X3). The dependent or outcome variable will be the performance score on the temporal bone dissection examination. Pre-study performance, or level of expertise in temporal bone surgical procedures, will be controlled for each resident by recording the following two factors:

1) performance on the pretest temporal bone examination
2) information from the American Board of Otolaryngology Operative Experience Report*

*Each resident is required by the American Board of Otolaryngology to record their operative experience. This will provide data on actual otologic surgical experience that can therefore be examined as a variable. These data are collected in a standardized format for all Otolaryngological training programs in the United States.

D.7.2 Procedure

D.7.2.1 Assignment

Before evaluation of performance, each resident will be randomly assigned to either, X1, X2 , or X3; thus each group will serve as a control or comparison group for the other two, in addition to controlling for other confounding factors.

D.7.2.2 Data Collection

To address for the fact that the more time the resident has spent on their training process, the more expertise he/she gains, the post-test examination will be performed at a fixed time of 2 weeks after the initial pre-test examination. We will then compare the posttest scores of each group, over the PGY3-5 levels. All pre and posttest temporal bones will be shipped to The Ohio State University College of Medicine and Public Health and stored until graded. Each bone will be graded as to the proficiency of dissection by 3 separate otologists based on the defined grading scale (see Welling Exam in Appendix). All otologist graders will be blinded to pre vs. posttest and X status of each bone. Additionally, to control for variability of cadaveric temporal bones, each bone will be categorized by an otologist grader regarding the level of difficulty of the inherent anatomy (1-easy,

2-moderate, 3-difficult). Expert graders will also make the same determination of for the virtual bones presented in the simulator.

D.7.3 Power and Statistical Analyses

Our evaluation seeks to determine the following:

- The predictive validity of the system
- The extent that the simulation is a useful learning tool and a good addition to traditional dissection

7.3.1 Objective trials

Initial exposure to the simulator will include an objective trial to familiarize the subject with the environment and with the general tasks that will be required in the study. All residents will be required before each pretest and posttest to execute an objective trial on the simulator. This objective trial employs a colored sphere inscribed inside a differently colored cube. The cube will be semitransparent so that as the user approaches the sphere, discoloration will be seen. Colors and transparencies will be the same as used in the subjective trials on the simulator. The subjects will be asked to use the various burrs to remove the outside cube and expose the sphere as precisely as possible. The time to complete the task and a measure of precision will be recorded. Descriptive statistics (mean, median, standard deviation, minimum and maximum value) will be reported.

As an exploratory analysis, we will compare the objective trial outcome variables (time to complete the task and precision) for the pre and post-tests, by using one-sample paired t-tests. If significant differences are found, we would fit a more complete model to learn more about the progress that residents made in the simulator and how confident they get as they are exposed more than once to this kind of experience. So, we will compute the difference in the outcome measurements (time and precision) between two weeks and baseline and perform one-way analyses of variance, using the two groups: a) no sessions and traditional vs. b) simulator, as the factor in the ANOVA.

7.3.2 Subjective trials

Each subject will then perform a cadaveric and a virtual temporal bone dissection to the best of their ability (pretest). The first group (X1), will receive no supplemental sessions. The second group (X2), will receive 6 hours of supplemental, un-mentored time in the traditional temporal bone dissection laboratory dissecting actual bones. A diary will be maintained documenting actual hours spent, number of bones dissected and any additional training received during this session to control for level of experience. The third group (X3), will receive 6 hours of supplemental un-mentored sessions as well, but on the temporal bone dissection simulator dissecting virtual bones. Lastly, each subject will perform an additional objective test followed by another cadaveric and virtual temporal bone dissection (posttest). Although not the focus of this proposal, this virtual temporal bone pre and post test data collection will help us look at the value of our system for testing and therefore provide insight into further development of a virtual testing environment that may give purely objective and quantifiable results.

The subjective trials will include performing the dissections on the simulator on a virtual bone or in the temporal bone lab on an actual bone. As an exploratory study, residents from years 3-5 will be recruited to use the simulator from The Ohio State University's Department of Otolaryngology. Note that only 3 residents per year enter the 5-year study program, making our study sample size too small to perform any statistical comparisons. With a sample size of 3 residents per group and an alpha level of 0.05, we could only detect an effect size of 3.1 with 80% power, which is too large to be clinically relevant. Therefore, only descriptive statistics and summaries will be reported for this study. The main purpose of this is to insure that every detail will be taken care for in the multi-institutional study.

In the multi-institutional study of the simulator, subjects will be randomized to any of the three groups: I) no sessions (XI), II) traditional temporal bone lab sessions (X2) and III) simulator temporal bone sessions (X3). Each resident will perform a temporal bone dissection and a simulator bone session at baseline and 2 weeks after being assigned to a group. With a sample size of 37 residents per group, we will have 80% power to detect an effect size of 0.67, between any two of the three different groups, using two-sided tests with a 0.05 significance level. This effect size corresponds to a difference in scores of 6 points and a standard deviation of 9 points in the Welling scale. Dr. Welling (pers com) indicated that the range of standard deviations that he has observed is 6.5 to 9 points, and that mean differences of 6 to 8 points are important from the clinical point of

view. Thus, the above sample size calculation is based on the most conservative scenario (using the smallest mean difference and the largest standard deviation). In addition, we think that the sample size of 37 per group is too conservative, since at least 30 centers already agreed on being part of the study. Multiplicity adjustments will be performed to adjust for multiple comparisons, if necessary.

The difference between the scores at two weeks and at baseline will be computed. The difference in scores between three groups will then be compared using a multiple linear regression model adjusting for potential confounders. Potential confounding factors such as hours spent studying, hours dissecting in the temporal bone lab, number of bones, level of bone difficulty (this is a categorical variable with three levels: easy, moderate and difficult) and number of assisted surgeries will be considered in this model. Also, all relevant two-way interactions will be included in the model. Those with an alpha level lesser or equal to 0.01 will be kept in the model, as well as the main effects involved in the interactions. When interpreting the model, more emphasis will be given to the existence of those significant interactions rather than the main effects involved in the significant interactions. Using the rule of thumb of ten observations per variable (Netter et al., 1996, Ch. 10), with this sample size we could include at most 11 variables in the model to study their significance.

Also, the correlation between Welling scores and the objective task outcome variables will be computed for the pre and post-tests. This will be important to determine if objective task outcome variables are good predictors of their performance in the simulator dissecting the virtual bone.

The proposed sample size was based on the assumption that parametric statistical methods will be used. But, sometimes the amount of missing data forces the investigator to use different statistical analyses than those proposed. Thus, if this is the case the appropriate non-parametric methods will be used in each case. In addition we will compare the amount of missing data in each group since that may be an indication of potential problems.

D.8 TimeLine

Phase I - Year 1
Initiate image acquisition from cadaveric specimens
Refine imaging protocols for improved structural acquisition
Establish web site for disseminating system updates, training and software
Establish ”Birds of a Feather” session at national meetings
Continue porting system to PC platform to Linux and Windows
Initiate development of dynamic, time-series, and multi resolution software
Complete Integration of acquired data into simulator
Complete software for capture of metrics and for use in local validation studies
Recruit and initiate local validation studies
Initial Reports

Phase II – Year 2
Continue refining image acquisition protocols including 4.7 T MR.
Continue structural data acquisition from cadaveric specimens
Establish method to acquire temporal bone registry data
Continue integration of data sets and refinement of realism
Continue development of dynamic, time-series, and multi-resolution software
Interim Report

Phase III – Year 3,4 & 5
Complete refining image acquisition protocols for all imaging modalities
Complete integration of temporal bone registry data
Continue ”Birds of a Feather” session at national meetings
Refine and complete dynamic, time-series, and multi-resolution software
Initiate and complete randomized multi-institutional study of the simulation
Final Report and Publications.