Anatomical Ground Truth: Reality or Illusion?

Russell Gullekson*

Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, New York, USA

ABSTRACT

Misinterpretation Reference ranges and statistical atlases are commonly used in medicine to assist physicians in determining what is abnormal. Originally, organ size and shape were confined to simple weight or linear measurements, including those taken from imaging, such as x-rays. Yet in this computational era, modern datasets are primarily developed using medical imaging. However, many factors can affect the data contained within these databases, and medical imaging techniques each have their own advantages, limitations and disadvantages. The three most common medical imaging methods (ultrasound, CT, and MRI) are discussed, and optical imaging methods are introduced.

Anatomical ground truth is the baseline by which the veracity of anatomic data collection methods can be measured. Logically, determination of anatomical ground truth would optimally rely on the direct physical measurement of human organs and structures. Obviously, widespread direct physical access to human organs is limited, and primarily consists of transplanted or pathological organs, including those that have experienced age-related changes. However, the necessity for anatomical ground truth still needs to be addressed as a foundation of both anatomic science, diagnostic imaging, and medicine itself. Discussion of what constitutes anatomical ground truth, and if it is even possible to achieve, is followed by consideration of what is required to achieve anatomical ground truth.

Keywords: Anatomical ground truth; Organ size; Reference range

***Correspondence to**: Russell Gullekson, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, New York, USA

Received: Jan 15, 2025, 2025; **Accepted**: Mar 07, 2025; **Published**: Mar 14, 2025

Citation: Gullekson R (2025) Anatomical Ground Truth: Reality or Illusion? J Anatomical Variation and Clinical Case Report 3:116. DOI: https://doi.org/10.61309/javccr.1000116

INTRODUCTION

Diversity exists in any biological system. Recognizing variation in anatomical structure is important in science, engineering, and especially medicine. Science and medicine are interested in the relationship between structure and function in general. In order to investigate structure and function in any biological system, the structure needs to be known with accuracy, and we need to know how that structure is related to function as well as existing factors that are impacting morphology at the time of study.

Perhaps more to the point, is how one answers what a structure *is*. The philosophy of "what is" underlies how something is named, in addition to how one perceives the 'reality' of the object. It's a fundamental issue for science across the board, whether looking at molecules, organ structure, or chemistry. There remains the need to establish "what is" the structure of interest, and decide among different ways of answering the question.

Ground truth is information obtained through empirical observation that is known to be both accurate and correct. Essentially, the term derives from the idea of a fundamental fact that is known to be authentic and verified, and therefore *true*. It is a requirement for development and analysis of any theoretical construct or design. In medicine, anatomical ground truth is generally assumed, based upon what empirical studies are available and in conjunction with established clinical application of medical imaging.

It has long been known that organ size and shape can reflect the health of the structure. Classically, organs and anatomic structures have been described linguistically [1], with the addition of some physical measurements, primarily in the later part of the twentieth century. Predominantly, organ weight [2-4], linear measurements [5-10], or volume [11,12] have been used to determine a normal range of variability. Reference ranges are a set of measurements in otherwise healthy individuals that is used by physicians to help determine abnormal findings. Statistical atlases are computational tools that can be used to compare reference ranges with patient-specific data. However, existing datasets are not universally accessible, and may be limited in the number of variables that they contain, and be based on small populations. Reference ranges are available from previous studies, but the population datasets are not often expanded once established. Currently, there is no standardization of such datasets. Reference ranges available from actual cadaveric organs are generally dated or from limited population samples, and have not been compiled into a single database.

Historically, reference ranges have been created using small sample populations, often with a focus on or restriction to males [13-15] and/or sample populations of limited geographical ancestry (e.g. Caucasian and/or African). In fact, considering the biases against population diversity and changes in genetic makeup resulting in an increase of individuals of blended or ambiguous phenotypic morphology, one can question whether reference ranges derived more than even a decade ago are clinically valid. With era and institutional biases inherent in reference ranges obtained throughout the twentieth century [16,17] combined with an increased admixture of genealogy due to immigration [18], increasing ease of travel, and relaxing social class and race restrictions, updated reference ranges are desperately needed [19]. With an evolving understanding of the existence and effects of sexual dimorphism and genotype on measured variables, reference populations have increased in diversity over time, with an increasing frequency of expanding observations to females [20] and individuals of non-Caucasian ancestry. However, available data on many groups of geographical ancestry is far less than that of other groups. It is also known that organ size regulation is a complex interplay of selective pressures. As such, organ size is not temporally static within an individual, but must exhibit a measure of plasticity [21,22]. This enhances the demand for reference ranges that investigate organ morphology changes over time, or at least in different age-deliminated populations.

There are many ways to look at organ structure, and novel methods for obtaining structure information are available. Therefore, anatomical ground truth must be known in order to properly assess the structure. Anatomical ground truth has varying definitions, depending on the method of acquiring data. Methods for measuring anatomical structure essentially fall into two categories: direct physical measurement and measurement from some form of imaging. There are logistic, technological, and perceptual hurdles to recognize when selecting a method of measurement. Classically, anatomic phenotypic descriptions are lacking in computational definition. Efforts to address this deficiency over the past decade-plus have had varying success, and constitute modern anatomical databases. However, there are not enough studies utilizing the direct measurement of organs in modern populations, and few recent studies have used actual physical measurement of explanted organs in comparison with less direct methods in order to establish the veracity of non-physical data acquisition.

In modern reference ranges and statistical atlases, medical imaging is usually considered to be ground truth data. Nevertheless, imaging is an electromagnetic representation of structure, not ultimate ground truth. Every radiological representation of a structure has sources for error, so establishing anatomical ground truth is critical in understanding the error. Therefore, when one considers the sources of possible error in a given imaging system, there still exists the need to corroborate the fidelity of medical imaging to the actual physicality of anatomical structures [23].

The arrival of the modern computerized era allows the computational description of organs through new, novel methods of analysis. Modern reference populations and statistical atlases are composed of computerized datasets that describe organs. These databases are built from data obtained primarily through medical imaging, which itself contains a wide variety of methods, resolutions, and accuracy. Medical imaging techniques offer a wide range of methods to study organ function in addition to morphology, and as such fulfills a critical role in diagnosing and treating disease. However, medical imaging has rarely been validated through comparison to measurement of the physical organs themselves, for obvious reasons. This lack of corroboration of electronic representations obtained through imaging with the actual anatomical structures themselves is a gap that can lead to large, possibly incorrect, assumptions that could have ultimately dire consequences.

Anatomical ground truth must satisfy specific requirements. It must provide the baseline information needed to utilize reference ranges in a clinical setting. It must also allow accurate interpretation of images and segmentation of structures from background tissues. Anatomical ground truth is crucial for developing and validating medical imaging, as well as image analysis algorithms and techniques. It also must provide a baseline reference for teaching medical imaging. Accepting any data as ground truth without satisfying all of these criteria must be conducted with the utmost caution and deliberation.

IS ANATOMICAL-GROUND TRUTH ACTUALLY POSSIBLE?

One cannot deny the usefulness of direct physical measurement in describing organ size and shape, and it can be used as a benchmark for the consideration of data obtained through other methods. It can ensure the accuracy and reliability of reference ranges, statistical atlases, and individual clinical diagnostic imaging by providing a baseline of physical reality. It can also provide a means by which to normalize measurements and characteristics across populations and studies, thereby establishing a standardized means to interpret other data. When used as a baseline, anatomical ground truth can provide a means by which to check quality of images and measurement data. In fact, anatomical ground truth cannot be completely understood or defined for practical use without direct physical measurement of organs as a baseline reference. As such, let’s assume that direct physical measurement of organs can be considered the ultimate anatomical ground truth.

Direct physical measurement of organs presents its own obvious challenges. Many organs are absolutely crucial for survival, and even when not, the lack can have an adverse physical and physiological effects on the individual. Therefore, direct physical measurement of organs must be limited to organs that are removed due to medical necessity or are obtained from cadavers. This is quite limiting, especially as organs removed from a living subject are most often pathological. Also, relatively few individuals donate their bodies for medical education and research. Combined with the limited resources of anatomical gift programs that receive cadavers for such purposes, it would behoove science and medicine to utilize every opportunity to obtain data from direct physical measurement of organs. Another opportunity for direct physical measurement of organs would be using organs harvested for transplantation, although measurement methods and time limitations for successful transplantation may preclude many such opportunities.

There is a need for a computational and analytical pipeline that defines anatomical structure in a mathematical way in order to properly compare direct physical measurement of organs to medical imaging. This would include a standardized list of measurements and how they are obtained. A centralized database of anatomical data would be essential for widespread use, and an easily accessible means by which to incorporate new data. The limitless possibilities of a publicly available repository of anatomical data would be of incalculable value.

Because anatomical variation is common and anomalies exist, anatomical ground truth must include such findings as part of normal anatomy. Anatomical ground truth cannot be considered reliable and encompassing without including common variations and anomalies as normal. In addition, there is no absolute with anatomic range, but statistically significant differences can be calibrated against diseased specimens, age groups, and levels of chronicity. Therefore, data on variations and anomalies must be included and noted in reference ranges and statistical atlases, which will allow for the identification and differentiation of anatomical variation from clinically relevant abnormal findings. Reference ranges that stratify disease progression would be extremely useful in treatment, and it is crucial that reference ranges for organs in diseased states also be established. It may also be that sub-ranges can be observed even within healthy populations. Modern-era reference ranges should also include gender-transitioned individuals and other underrepresented groups in order to achieve the clearest pictures of anatomical variation, disease states, and public health.

When looking at anatomical structures, the answer to “What is ground truth?” is defined by what the purpose or question is, as well as the methodology used to answer the question. This would define how one would look at an anatomical structure. For example, there may be a need in an ER to determine the size of a patient’s kidney. Ultrasound is usually available, but what should that information be compared to? What reference range should be used? Ultrasound is a sonic representation; it’s not ‘truth’, yet it is what is available.

Ultimately, anatomical ground truth should be possible, although we must first decide exactly what that is, being measurement of a physical object directly or of non-physical representations. Direct physical measurement of organs is certainly possible, but it is time-consuming and lacks funding, and possibly interest as a result. The limited availability of organs to measure complicates the issue, as data would be procured slowly, and will always be insufficient in quantity. However, impracticality doesn't outweigh the need.

MEDICAL IMAGING AND ORGAN SIZE

Medical imaging has been used as a surrogate for organ measurement in living patients for well over a century, beginning with x-rays. A plethora of studies have established the clinical relevance and application of medical imaging, which is used in diagnosis and treatment, as well as to address biological queries. In modern medicine, ultrasound, CT, and MRI are the primary modalities used to acquire medical images of soft tissue structures, although other imaging modalities are available. Most others are used to obtain specific information that is not available through other techniques. It is important to understand that each imaging modality possesses its own set of uses, limitations, and sources of error. All imaging technologies have varying resolution capabilities even within the method, based on machine power, age of the technology, hardware, and supporting software. Combination technologies, such as PET/CT, increase the capabilities of each technology by producing more information with greater detail. Imaging potential is always improving, with advances in power/output and processing algorithms, although there are risks in addition to benefits for every imaging method [24].

ASSUMPTIONS, LIMITATIONS, AND SOURCES OF ERROR

Accuracy describes how close the scan reflects the actual measurements of the scanned object. Resolution describes the level of detail that a scanner can capture from an object. Resolution capability varies based on hardware, but the object being scanned and the purpose of the scan determines the necessary scan resolution.

Medical imaging techniques have a number of common limitations. Most are costly, requiring (often extremely) expensive equipment and infrastructure, such as weight-bearing requirements. They also necessitate operation by highly trained, skilled staff. Some imaging modalities present a health risk due to ionizing radiation, reactions to injected contrast media, or the need for a patient to remain in a specific position. Resolution varies, depending on the age and quality of the equipment, operator knowledge, training, and skill, and the software used to process the image. Measurement equipment needs to be regularly and properly calibrated, and software and computer screens that utilize color for structure identification and demarcation generally have subjective settings. No imaging modality is best for all organs, structures, and tissues, so the strengths and weaknesses of each imaging technique must be evaluated on a case-by-case basis for appropriateness in order to establish the best available imaging option.

Most imaging modalities require that the patient remain immobile, as movement directly affects the image. In addition, within each imaging modality, individual organs or structures must be manually deliminated from surrounding organs and tissues by a (preferably) qualified radiologist, a process called segmentation. Segmentation takes a significant amount of time, especially in 3D modalities, and is essentially performed freehand [25]. This introduction of possible human error is important to recognize. Much effort over the past decade plus has been put into harnessing computers to use training populations to educate artificial intelligence (AI) to auto-segment organs from radiological scans, with varying success. However, all attempts must ultimately still be compared to manual segmentation to assess accuracy [26].

Computational anatomy is the mathematical study of anatomy in which quantitative analysis and modelling of the variation in anatomical shape is performed. In this field, template organs are compared to individuals to discover differences that may indicate disease. Strides have been made in recent years in this field, although it has only existed in this century. However, it is still a developing science. As with segmentation, the greater the amount of ground truth data available, the greater the performance of computational systems.

A review of the literature shows that very few studies have compared measurements from imaging modalities to direct physical measurement of the same specimens, and the studies available generally provide little description of how the measurements were performed. Available papers also generally focus on specific anatomic structures (e.g. size of tumors or other pathologies, various dimensions of bones and joints, or organ substructures), and not whole organs. The literature also exhibits a lack of information on how modern imaging technologies were developed. It may be assumed that as new imaging methods were developed, they were calibrated using phantoms of known shape and size, and images produced were compared in accuracy and resolution to other, pre-existing imaging methods. This lack of comparison between medical imaging direct physical quantification of unique shapes highlights an incredible assumption, that representations of anatomic structures are accurate, without the need for correlation testing against the physical object itself.

Additionally, background or signal noise is an ever-present reality in imaging, and affects the number of photons detected, affecting the graininess of the image. Scattered secondary radiation, image power, slice thickness, patient body size and composition, and other factors affecting the number of photons reaching detectors can all produce noise [27]. Do radiologists actually know the signal noise vis-à-vis ground truth? They should.

IMAGING MODALITIES

Ultrasound: Diagnostic ultrasound is a common medical imaging technique that uses sound waves to produce an image of the desired structure, and is the standard for many diagnostic and treatment procedures in medicine. It offers the benefit of being safe (in that it emits no ionizing radiation), relatively non-invasive, portable, simple, fast, and relatively inexpensive compared to most other medical imaging modalities. It can also benefit from the ability of the subject to change position to assist the operator in acquiring an optimal image. 3D measurement of distance and volume is more accurate than 2D [28], although both versions are considered to be accurate enough for clinical use. However, ultrasound is generally considered to be less accurate than either CT or MRI.

Ultrasound examinations are limited in the portion of the body that can be imaged during an examination. Additionally, the sound waves are blocked by higher-density tissues such as bone, and deeper structures require a lower frequency, which can result in a decrease in image resolution. Image artifacts are common, although multi-beam technology decreases the incidence of artifacts. Some tissues and objects (e.g. nerves, tendons, and needles) reflect the ultrasound beam (anisotropy). Ultrasound imaging also requires training and skill. Ultrasound is limited by tissue depth, and it can be difficult to distinguish between different tissues. Ultrasound image resolution can be affected by many factors, including tissue density, the presence of air or bone, and the angle of the sound waves. Many things can result in error in ultrasound imaging, including equipment limitations, and operator skill, knowledge, and technique [27,29,30]. Ultrasound waves can also be attenuated or scattered by passing through tissues with varying properties. There have also been reports of changes in hormone levels after diagnostic ultrasound exposure and concerns involving effects on the fetus *in utero*, including on DNA [24,31,32], although these effects are disputed [33-35].

Despite the limitations, disadvantages, and opportunities for error in ultrasound imaging, it remains a valued technique. The considerations of patient safety, low cost, and ease of use are considerable benefits that contribute to this being a preferential method of imaging, when suitable. Ultrasound imaging is familiar to many, may be considered by patients to be less intimidating than the giant machines needed for other medical imaging (e.g. CT or MRI), and its portability contributes to the longevity and persistence of this technology.

Computed tomography: Computed tomography (CT) scanning is a technology in which digital x-ray images are used to create cross-sectional representations, or “slices”. Successive image layers are gathered together and ‘stacked’ to produce a 3D representation The resolution of a CT image is dependent upon the power of the x-rays being produced, the diameter and number of detector elements and their distance from the x-ray source, and the size of the x-ray focal site [36].

Computed tomography provides a fast acquisition of images, and can image an entire body or a region of interest at the same time. Unlike with standard x-rays, overlapping structures are prevented by the three-dimensional image being produced by x-rays taken at different angles. Individual slices can be viewed independently or as part of the 3D reconstruction of the body. CT scans can be used to gauge the size, shape, density, or even texture of an organ or structure. It can also be used to establish the exact location of structures and abnormalities, as it is having better spatial resolution than either ultrasound or Magnetic resonance imaging. Clinically, CT is often used to diagnose tumors, bleeding, or fractures.

Unfortunately, x-rays produce ionizing radiation, which is a health risk that increases the chance of cancer occurring. Although the higher dose of x-rays produced by CT than standard x-ray still results in a small cancer risk, it does exist [24,37]. This ionizing radiation may be especially harmful to children and pregnant women. In addition, Contrast agents, which are substances that prevent the passage of x-rays through it, allowing for better resolution and tissue differentiation, are often used. However, there is also a risk of allergic reactions to contrast agents, and in rare cases kidney damage has occurred [38]. Therefore, the benefit of a CT scan and the use of contrast agents must be considered against the risks.

CT scanners are large, heavy machines that require extensive physical infrastructure, including reinforced floors, ample power, a separate room for the CT computer, and walls that are impenetrable to ionizing radiation. Additionally, the aperture size of the scanner and weight limit of the patient bed can limit the ability to perform CT with some patients. While mobile CT scanners exist, they are uncommon. Therefore, CT systems are very costly, with new scanners ranging from around $300,000 to $2 million, depending on the number of slices the machine can produce in a single rotation. And that is in addition to the costs of supporting infrastructure, including computers and storage for large DICOM file sets. Additionally, training, certification, and radiation safety must be factored into operational costs. Due to the enormous expenditures involved with CT scanning, the cost of individual scans can range anywhere from hundreds to thousands of dollars, depending on the type of facility, geographic region, quality of the machine, and anatomic extent of the scan. CT technology may not be available in an area due to the costs involved. Also, as there is rarely more than one CT scanner at a site due to cost, the technology may be unavailable if the only accessible scanner is non-functioning. Like any high-tech machine, CT scanners are vulnerable to breakdown and malfunction, sometimes resulting in lengthy repair periods.

Numerous factors may introduce a measure of error into the accuracy and resolution of a CT scan. Metallic implants can scatter the x-rays, and patient motion may create registration artifacts such as blurring, streaking, or shading. Beam hardening occurs when x-rays pass through an object that acts to ‘filter out’ lower-energy photons, and also results in image errors. Partial volume effects result when multiple tissues are present in a single voxel. This results in tissue boundaries becoming blurry. Misalignments of key components such as the x-ray focal location, detectors, and rotation stage can result in image errors [39]. Background noise in the CT system, most commonly from either ‘quantum’ noise associated with the number of photons detected or ‘electronic’ noise from the detector system, is also possible [27,40].

CT remains a common imaging modality in medicine, due to its balance of speed, resolution, and cost. The accuracy and resolution available make it a desirable choice when ultrasound isn’t enough. However, the production of ionizing radiation remains a concern that must be considered before selecting CT as the appropriate means of acquiring information for diagnosis.

Magnetic Resonance Imaging: Magnetic resonance imaging (MRI) is a technology that uses ultra-strong magnets to align the axes of hydrogen molecules in the body, creating a vector that is detectable by the scanner [41]. MRI is commonly used to image nervous tissue, ligaments, and tendons, as it produces more detailed images of such structures than CT scans. This is because a better tissue contrast is possible, providing better differentiation between adipose, water, and other soft tissues. Since MRI does not produce ionizing radiation, it is safer than CT. However, the strong magnets are never turned off, and will attract metallic objects of almost any size, therefore the machine room must be kept clear of metallic objects at all times. Injury from objects acting as projectiles has occurred.

The magnetic fields also create a safety hazard for people with implants (e.g. pacemaker or artificial hip joint), external medical devices (including prosthetics or braces) that have metal components [41], or accessory medical devices (such as IV stand, oxygen tank, or external defibrillator). The magnetic fields of the machine can cause unwanted movement of the devices due to the pull-on metallic materials, and the device or tissue could become heated by the radio waves [42].

Additionally, the changing of magnetic fields causes a loud knocking sound that may require ear protection for the patient. Lengthy exposure to radiofrequency energy can result in the generation of heat within the body, especially during lengthy examinations, potentially causing first or second-degree burns. Contrast agents, as with those used with CT, also pose similar risks [42]. The technology also requires patients to remain motionless, lest image artifacts be observed. This can be especially difficult for patients in pain or who encounter a claustrophobia-like experience inside the small machine aperture.

MRI machines are expensive, with new machine running as much as $3 million. As with CT, MRI requires a significant amount of supporting infrastructure. The machine is very heavy, requiring reinforced flooring, and all walls, floors, and ceilings must be shielded against the magnetic field. MRI also has data storage, computer technology, training and certification requirements. In fact, MRI requirements result in almost double the cost of CT, often resulting in delays based on the need for pre-approval from medical insurance carriers.

There are also numerous possible sources of error in MRI imaging, including issues with the magnetic field or interference with the radio waves that can also affect the resulting resolution. Aliasing is an imaging error in which imaging parameter mapping errors can occur due to normal noise and aliasing artifacts [43], and chemical shifts between the resonance frequencies of adipose and water can produce imaging artifacts [44,45]. In addition, the presence of metallic objects may decrease the quality of the image obtained.

Optical Imaging: Other modalities exist for acquiring surface renderings of organs. The equipment necessary for these modalities is easily obtainable, has a smaller footprint (being able to be placed on a table or counter), and costs significantly less than medical imaging. Although these methods require internal structures be explanted or otherwise revealed to the naked eye before use, these techniques may well be of use in establishing anatomical ground truth. Optical surface scanning, while incapable of imaging organ structure beneath the surface, can provide significant information on the size, shape, and even pathology of an organ. The external surfaces of kidneys, for example, can exhibit a lot of information about what's going on in the cortex, and therefore organ function status.

There are clinical situations where organs are removed and transplanted into another person, or removed due to a pathological necessity, but a healthy organ isn't going to be cut open or put in an MR or CT machine. Ultrasound could be performed on an explanted organ, but the volume of the kidney could be calculated fairly easily and quickly with photogrammetry or some form of optical surface scanning.

Laser scanning involves a scanner with laser(s), computer, and camera(s) to capture the structure surface at the point of contact with the laser to create a computational model for analysis or 3D printing. Laser scanning is limited to the surface of the structure, and obtained image quality can be affected by a number of factors, including lighting, reflectiveness of material or surface moisture, and complexity of the surface [46]. Some 3D laser scanners also use photogrammetric markers to increase the accuracy of the scan [47]. Although historically conducted using 2D film photography to measure objects, modern photogrammetry requires camera(s) to capture the object surface using digital photography. Using software to 'stitch' the photographic series together, photogrammetry can provide a similar 3D computational model.

DISCUSSION

In considering which imaging modality is closer to anatomical ground truth, or can best represent it, we must decide what the specific goal of the choice is. If it is to measure something, then whichever technology gives the best resolution for that structure would obviously be the best choice. If the question is more related to physiology or structural abnormality or pathology, then a different method may be more suitable. Additionally, when used with live patients, safety must be considered as well as available resources and cost effectiveness. However, the lack of studies providing direct physical validation of medical imaging logically suggest caution in assuming that any imaging modality accurately predicts or represents anatomical ground truth.

I submit that anatomical ground truth is not only possible, it is necessary. Direct physical measurement of anatomical structures is crucial but poorly represented data in reference ranges, and needs to be expanded. It may be that direct physical measurement of anatomical structure is the most 'true' of all measurement methods, in that there is little to affect the reality of obtained values. As such, it should be conducted when possible as a baseline by which other methods may be assessed for accuracy. The rarity of specimens and difficulty in obtaining this data only illuminates its importance. However, direct physical measurement data may not be readily available for clinical decision making, or unobtainable for specific research, in which case medical imaging is the obvious surrogate. Until we expand our knowledge of anatomical variation by examining physical structure directly, we will remain unsure of the true accuracy of medical imaging, leaving its use full of assumptions and possible misinterpretations. Therefore, direct physical measurement of anatomical structures needs to not only be conducted, it also needs to be considered as a primary method of obtaining anatomical morphological data.

When comparing the different methods of measuring & analyzing structure, one must consider many elements, including accessibility, education, cost, and practicality. Technologies and methodologies that are not widely available may provide valuable information, but be difficult to gain access to, be cost prohibitive, or be poorly understood. The education requirements for understanding the benefits and utility of a specific method may necessitate additional personnel, equipment, or other resources. Various methodologies may incur additional costs in infrastructure (e.g. building, energy, data transmission and storage needs), equipment and/or supplies, education, training, and certification, maintenance, etc. The practicality of a method must include other considerations, such as space needs, skill, speed of data acquisition, availability of data storage, proprietary software and innumerable additional factors. All of these considerations must be accounted for when deciding among data acquisition methods available for research or clinical application. The tradeoffs of each available system must obviously be examined against its benefits and disadvantages. This will lead to considering how to address current gaps and limitations in available techniques and technologies. While it may be possible to arrange the usage of multiple systems for a research project or clinical case, allowing additional methods to provide information lacking from others, multimodal investigation is often impractical, time-consuming, and cost prohibitive. In the modern era of privatized medicine and limited research funding, utilizing multiple methods may not be a realistic option.

There is an obvious need for anatomical ground truth data. As well-defined reference range populations advance the science of anatomy, it also furthers its application in medicine, by providing more accurate and realistic measures for what is normal. This can impact population health overall, as reference ranges are needed to manage the health of all demographic sub-groups. Theoretically, the predictive value of reference ranges and statistical atlases may be correlated with histopathology and clinical test values to create a more detailed picture of both health and disease states. Without more precise stratification of population data, population matching will always be left with assumptions about included subjects.

Expanded anatomical ground truth knowledge could enhance the efficiency of pathology. If anatomical data was both more readily available and of increased quantity, individual patient anatomical data (such as size, shape, or weight) could be compared to reference ranges more reliably and effectively, especially as a fast indicator of the possible presence of disease. For example, physicians might be more informed as to when further investigation was necessary, or pathologists knowing that a slide needed to be made from a tissue sample. Increased anatomical data could also contribute toward the development of ‘hands-free’ medical procedures such as virtual autopsies.

It may be that the reality of anatomical ground truth requires a multimodal, analytic point of view to create the sharpest understanding of the relationships between anatomical structure, disease, and public health. Each method of data acquisition, be it medical imaging or direct physical measurement, has value as well as limitations. Together, all methods can provide a more complete picture, and that is the ultimate goal. Medical imaging still requires further corroboration by comparison to direct physical measurement in order to validate its representations of the complex shapes of living anatomy.

REFERENCES

1. Standring S (2016) Gray's Anatomy: The anatomical basis of clinical practice. 41st ed, New York: Elsevier Limited.
2. Rao UV, Wagner HN (1972) Normal weights of human organs. Radiology 102: 337-339.
3. de la Grandmaison GL, Clairand I, Durigon M (2001) Organ weight in 684 adult autopsies: new tables for a Caucasoid population. Forensic Science International 119: 149-154.
4. Yang T, Kong J, Shao X, Meng Z, Liang P, et al. (2024) A statistical study of postmortem heart weight in Chinese adults. Forensic Science International 354: 111912.
5. Capaccioli L, Stecco A, Vanzi E, Brizzi E (2000) Ultrasonographic study on the growth and dimensions of healthy children and adults’ organs. Italian journal of anatomy and embryology 105: 1-50.
6. Duyar I, Pelin C (2003) Body height estimation based on tibia length in different stature groups. Am J Phys Anthropol 122: 23-27.
7. Kanitkar M (2012) Kidney size--what is normal? Indian pediatrics 49: 523.
8. Gholamzadeh S, Zarenezhad M, Montazeri M, Zareikordshooli M, Sadeghi G, et al. (2017) Statistical Analysis of Organ Morphometric Parameters and Weights in South Iranian Adult Autopsies. Medicine 96: e6447.
9. van der Sande NG, Blankestijn PJ, Leiner T, van der Graaf Y, de Borst GJ, et al. (2017) High ratios of kidney function to kidney size are related to mortality and kidney function decline in high-risk patients. European journal of preventive cardiology 24: 926-933.
10. Khan SA, Yasmeen S, Adel H, Adil SO, Huda F, et al. (2018) Sonographic Evaluation of Normal Liver, Spleen, and Renal Parameters in Adult Population: A Multicenter Study. JCPSP 28: 834-839.
11. Geraghty EM, Boone JM, McGahan JP, Jain K (2004) Normal organ volume assessment from abdominal CT. Abdom Imaging 29: 482-490.
12. Harbaugh CM, Shlykov MA, Tsuchida RE, Holcombe SA, Hirschl J, et al. (2015)

Morphometric analysis of abdominal organs and rib cage: Implication for risk of solid organ injuries in children. J Trauma Acute Care Surg 78: 1129-1133.
13. Zakiniaeiz Y, Cosgrove KP, Potenza MN, Mazure CM (2016) Balance of the Sexes: Addressing Sex Differences in Preclinical Research. Yale J Biol Med 89: 255-259.
14. Clayton JA (2018) Applying the new SABV (sex as a biological variable) policy to research and clinical care. Physiol Behav 187: 2-5.
15. Burrowes KS, Ruppage M, Lowry A, Zhao D (2023) Sex matters: the frequently overlooked importance of considering sex in computational models. Front Physiol 14: 1186646.
16. Gilpin NW, Taffe MA (2021) Toward an Anti-Racist Approach to Biomedical and Neuroscience Research. J Neurosci 41: 8669-8672.
17. Michael A, Bengtson J, Blatt S (2021) Genes, Race, Ancestry, and Identity in Forensic Anthropology: Historical Perspectives and Contemporary Concerns. Forensic Genomics 1: 41-46.
18. Quintyn C (2010) The Existence or Non-existence of Race: Forensic Anthropology, Population Admixture and the Future of Racial Classification in the U.S. New York: Teneo Press.
19. Muthusami P, Ananthakrishnan R, Santosh P (2014) Need for a nomogram of renal sizes in the Indian population-findings from a single centre sonographic study. The Indian journal of medical research 139: 686-693.
20. Arnegard ME, Whitten LA, Hunter C, Clayton JA (2020) Sex as a Biological Variable: A 5-Year Progress Report and Call to Action. J Womens Health 29: 858-864.
21. Spencer RP (1996) Organ/body weight loss with aging: evidence for co-ordinated involution. Med Hypotheses 46: 59-62.
22. Penzo-Méndez AI, Stanger BZ (2015) Organ-Size Regulation in Mammals. Cold Spring Harb Perspect Biol 7: a019240.
23. Alkemade A, Groot JM, Forstmann BU (2018) Do We Need a Human post mortem Whole-Brain Anatomical Ground Truth in in vivo Magnetic Resonance Imaging? Front Neuroanat 12: 110.
24. Hussain S, Mubeen I, Ullah N, Shah S, Khan BA, et al. (2022) Modern Diagnostic Imaging Technique Applications and Risk Factors in the Medical Field: A Review. BioMed research international 2022: 5164970.
25. Torres HR, Queiros S, Morais P, Oliveira B, Fonseca JC, et al. (2018) Kidney segmentation in ultrasound, magnetic resonance and computed tomography images: A systematic review. Computer methods and programs in biomedicine 157: 49-67.
26. Sistaninejhad B, Rasi H, Nayeri P (2023) A Review Paper about Deep Learning for Medical Image Analysis. Comput Math Methods Med 2023: 7091301.
27. Morin R, Mahesh M. 2018. Role of Noise in Medical Imaging. Journal of the American College of Radiology 15:1309.
28. Riccabona M, Nelson TR, Pretorius DH. 1996. Three-dimensional ultrasound: accuracy of distance and volume measurements. Ultrasound in obstetrics & gynecology: the official journal of the International Society of Ultrasound in Obstetrics and Gynecology 7:429-434.
29. Walas MK, Skoczylas K, Gierbliński I. 2012. Errors and mistakes in the ultrasound diagnostics of the liver, gallbladder and bile ducts. J Ultrason 12:446-462.
30. Di Serafino M, Iacobellis F, Schillirò ML, D'Auria D, Verde F, Grimaldi D, Dell'Aversano Orabona G, Caruso M, Sabatino V, Rinaldo C, Guerriero P, Cantisani V, Vallone G, Romano L. 2022. Common and Uncommon Errors in Emergency Ultrasound. Diagnostics (Basel) 12.
31. Beech B. 2002. Ultrasound revisited. Midwifery Today Int Midwife:28; author reply 28-29.
32. Beech BL. 2013. Ultrasound: prematurity and potential risks. Midwifery Today Int Midwife:26-28, 67.
33. Bromberg IL, Zielinski N, Dassin R, Fisher S. 1990. Hormonal effects of ultrasound. Epidemiology 1:330-331.
34. Ziskin MC. 1999. Intrauterine effects of ultrasound: human epidemiology. Teratology 59:252-260.
35. Abramowicz JS, Fowlkes JB, Skelly AC, Stratmeyer ME, Ziskin MC, et al. (2008) Conclusions regarding epidemiology for obstetric ultrasound. J Ultrasound Med 27: 637-644.
36. Uldin T. 2017. Virtual anthropology - a brief review of the literature and history of computed tomography. Forensic Sci Res 2:165-173.
37. Puchalski AL, Magill C. 2018. Imaging Gently. Emerg Med Clin North Am 36:349-368.
38. Andreucci M, Faga T, Pisani A, Sabbatini M, Michael A, et al. (2014) Acute Kidney

Injury by Radiographic Contrast Media: Pathogenesis and Prevention. BioMed research international 1-21.
39. Ferrucci M, Heřmánek P, Ametova E, Carmignato S, Dewulf W (2018) Measurement of the X-ray computed tomography instrument geometry by minimization of reprojection errors-implementation on simulated data. Precision Engineering 54: 7-20.
40. Yu Z, Leng S, Kappler S, Hahn K, Li Z, et al. (2016) Noise performance of low-dose CT: comparison between an energy integrating detector and a photon counting detector using a whole-body research photon counting CT scanner. J Med Imaging 3: 043503.
41. Berger A (2002) Magnetic resonance imaging. BMJ 324: 35.
42. Ghadimi M, Sapra A (2023) Magnetic Resonance Imaging Contraindications. StatPearls, StatPearls Publishing.
43. Kara D, Fan M, Hamilton J, Griswold M, Seiberlich N, et al. (2019) Parameter map error due to normal noise and aliasing artifacts in MR fingerprinting. Magn Reson Med 81: 3108-3123.
44. Graves MJ, Mitchell DG (2013) Body MRI artifacts in clinical practice: A physicist's and radiologist's perspective. J Magn Reson Imaging 38: 269-287.
45. Huang SY, Seethamraju RT, Patel P, Hahn PF, Kirsch JE, et al. (2015) Body MR Imaging: Artifacts, k-Space, and Solutions. Radiographics 35: 1439-1460.
46. Lemeš S, Zaimović-Uzunović N (2009) Study of ambient light influence on laser 3D scanning. In: Proceedings of the 7th international conference on industrial tools and material processing technologies 327-330.
47. Cherdo L (2022) 3D scanner accuracy and resolution: The basics. Bayonne, France: ANIWAA. Accessed October 10, 2023.